Multiscale Molecular Dynamics Studies of Peptide Aggregation and Protein-Ligand Association

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Andrea Magno
aus
Italien

Promotionskomitee

Prof. Dr. Amedeo Caflisch (Vorsitz)
Prof. Dr. Ben Schuler

Zürich 2013

Al mio “amico fragile” François

Summary

The focus of the present thesis is the use of different, multiscale molecular dynamics (MD) approaches to investigate the microscopic mechanisms underlying complex molecular processes, namely peptide aggregation and protein-ligand association. A first project deals with the employment and extension of a phenomenological coarse-grained model developed by Dr. Riccardo Pellarin and Prof. Dr. Amedeo Caflisch in order to assess the influence of macromolecular crowding on the aggregation kinetics of a model peptide. MD simulations show that the intrinsic amyloidogenic propensity of the peptide determines the extent to which the presence of large and inert cosolutes (i.e., crowders) accelerate the peptide self-assembly. In a second project, in collaboration with the group of Prof. E. Gazit at the University of Tel Aviv, we performed MD simulations to obtain microscopic details on the structure of the early aggregates formed by the amino acid phenylalanine. Our results show that this monopeptide is able to aggregate into ordered assemblies and confirm the data acquired with electron diffraction experiments. We further studied how carnosine, a putative small inhibitor of amyloid aggregation, interacts with the amyloid β monomer (Aβ) and affects its conformational ensemble and aggregation propensity. In agreement with the various experiments performed at the University of Udine and University of Catania, our simulations reveal that carnosine establishes transient and non-specific interactions with the Aβ monomer, without modifying its structural features. However, a reduced tendency to interpeptide hydrogen bonds was observed for Aβ peptide segments in presence of carnosine. Finally, full atomistic MD simulations in conjunction with an explicit representation of the aqueous environment were used to shed light on the intrinsic flexibility of the binding sites of 20 bromodomains, protein modules which specifically recognize acetylated lysine, and to unveil the mechanism of acetyl-lysine binding to the bromodomain TAF1(2). We observed a large variability in the orientation of several conserved residues in the bromodomains binding groove, which complete the static picture provided by crystallographic experiments. Moreover, our computational approach proved able to reproduce the acetyl-lysine binding pose observed in the X-ray structures and unmasked the existence of a new stable binding conformation, in which the acetyl-lysine side chain is more buried in the binding pocket and displaces some of the structured water molecules at the bottom of the binding cleft.

Zusammenfassung

Der Fokus der vorliegenden Thesis liegt auf dem Einsatz von verschiedenen, multiskalen Moleküldynamik- (MD-) Methoden, um die mikroskopischen Mechanismen, welche molekularen Prozessen zu Grunde liegen, zu untersuchen. Speziell wurden die Peptid-Aggregation und die Protein-Ligand Bindung betrachtet. Ein erstes Projekt handelt von der Anwendung und Erweiterung eines phänomenologischen grobkörnigen Modells, das von Dr. Riccardo Pellarin und Prof. Amedeo Caflisch entwickelt wurde. Mit Hilfe dessen kann man untersuchen, wie makromolekulare Verdrängung die Aggregationskinetik eines Modellpeptids beeinflusst. MD-Simulationen zeigen, dass die intrinsische Amyloid-Tendenz des Peptids bestimmt, inwiefern die Präsenz grosser und inerter Kosolute (bzw. Crowders) die Peptid-Aggregation beschleunigt. Für ein erstes Projekt, eine Kollaboration mit der Gruppe von Prof. E. Gazit an der Universität Tel Aviv, haben wir MD-Simulationen ausgeführt, um mikroskopische Details über die Struktur der Frühaggregate der Aminosäure Phenylalanin zu erhalten. Gemäss unserer Ergebnisse hat dieses Monopeptid die Fähigkeit, geordnete Aggregate zu bilden. Ausserdem bestätigen unsere Resultate die Daten, welche mit Elektronen-Diffraktion gefunden wurden. Weiter haben wir erforscht, wie Carnosin, ein mutmasslicher Hemmstoff der Amyloid-Aggregation, mit dem Amyloid-β (Aβ) Monomer wechselwirkt und wie es das konformative Ensemble und die Amyloid-Tendenz von Aβ beinflusst. In Übereinstimmung mit den verschiedenen Experimenten, die an der Universität Udine und Universität Catania durchgeführt wurden, enthüllen unsere Simulationen, dass Carnosin mit dem Aβ Monomer transiente und unspezifische Wechselwirkungen formt, ohne seine strukturellen Eigenschaften zu

verändern. Trotzdem wurde für Aβ Peptid-Segmente im Beisein von Carnosin eine reduzierte Tendenz zu interpeptidischen Wasserstoffbrücken beobachtet. Zuletzt wurden vollständig atomistische MD-Simulationen zusammen mit einer expliziten Darstellung der wässigren Umgebung benutzt, um die intrinsische Flexibilität der Bindungsstelle von 20 Bromodomänen, Proteinmodule, welche Acetyl-Lysin spezifisch erkennen, ans Licht zu bringen und den Mechanismus der Acetyl-Lysin Bindung an die Bromodomäne TAF1(2) aufzudecken. Wir beobachteten, dass verschiedene konservierte Residuen in der Bindungsstelle von Bromodomänen mehrere verschiedene Orientierungen einnehmen können. Das kompletiert das statische Bild, welches kristallographische Versuche schaffen. Ausserdem wurde gezeigt, dass unsere computergestützte Herangehensweise die Fähigkeit besitzt, die in Kristall-Strukturen gefundene Bindungspose des Acetyl-Lysins zu reproduzieren. Zusätzlich enthüllten die MD-Simulationen eine neue stabile Bindungspose, in der die Seitenkette des Acetyl-Lysins tiefer in der Bindungsstelle liegt und einige strukturierte Wassermoleküle in der Bindungsstelle ersetzt.

Contents

Chapter 1

Introduction

Computer simulations are nowadays an established research tool used in a variety of scientific fields, from condensed matter physics and astrophysics to meteorology and economics. In biology, *in silico* experiments are routinely performed to complement and extend *in vivo* and *in vitro* investigations which are often limited in time and spatial resolution. They are particularly useful to study the kinetic properties, as well as the thermodynamic stability and structure of the different states encountered within a molecular system during a specific process, i.e., protein folding and aggregation, ligand binding and enzymatic catalysis. The real experiment is modeled and replicated in the computer and some observables are measured and compared with known data. If an agreement between the model and the experiment is observed, the simulation can be used to make predictions to be further tested experimentally. Therefore, computational studies represent a sort of “middle way” between theoretical models and experiments [1]: on one hand they allow to test the validity of theoretical predictions on a certain phenomenon, on the other hand they can shed light on details that would be otherwise unaccessible to experimental techniques.

1.1 Molecular dynamics simulations

In 1957 Alder and Wainwright published the first molecular dynamics (MD) study of a system of hard spheres [2]. This computational method reproduced the evolution in time of the investigated system by solving the equation of motion for its constituents. Today, there are essentially two types of MD simulations [3]:

- *ab initio* simulations, in which the degrees of freedom of electrons and atomic nuclei are explicitly taken into account and, therefore, the laws of quantum mechanics are used;
- molecular mechanics simulations, which are based on classical mechanics.

In the latter representation, a molecular system is represented as a collection of point masses moving in a conservative field, often called *force field*, which dictates the structure-energy relationship for the studied system. Several force fields for proteins and nucleic acids are currently available, such as CHARMM [4, 5], AMBER [6], OPLS [7] and GROMOS [8]. In general, the force field is given by the following pairwise function:

$$E(\mathbf{r}) = E_{cov}(\mathbf{r}) + E_{noncov}(\mathbf{r}) \tag{1.1}$$

where E_{cov} and E_{noncov} are the covalent and non-covalent energy terms, respectively and $\mathbf{r}$ is the $3N$ set of atomic coordinates. We will focus hereafter on the CHARMM force field. The covalent term is given by the sum over the energies of bonds b, angles θ, dihedral angles ϕ and improper dihedral angles ω:

$$\begin{aligned} E_{cov}(\mathbf{r}) &= E_b + E_\theta + E_\phi + E_\omega = \\ &= \sum_b k_b(b-b_0)^2 + \sum_\theta k_\theta(\theta-\theta_0)^2 + \\ &\sum_\phi k_\phi[1+cos(n\phi-\delta)] + \sum_\omega k_\omega(\omega-\omega_0)^2 \end{aligned} \tag{1.2}$$

where k_b, k_θ, k_ϕ and k_ω are the force constants, b_0, θ_0 and ω_0 represent the minimum of the corresponding harmonic potential, and n and δ govern the position and the multiplicity of the minimum of the dihedral periodic potential.

The non-covalent term in 1.1 is given by the sum of Lennard-Jones (LJ) and Coulombic (C) potentials:

$$\begin{aligned} E_{noncov}(\mathbf{r}) &= E_{LJ} + E_C = \\ &= \sum_{i \neq j} \epsilon_{ij} \left[\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{6} \right] + \sum_{i \neq j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \end{aligned} \tag{1.3}$$

where q_i stands for the partial charge of atom i, ϵ_0 is the vacuum permittivity, r_{ij}^{min} is the equilibrium separation between atoms i and j and ϵ_{ij} is their equilibrium LJ interaction energy. The Lennard-Jones potential mimics the van der Waals interactions arising from the impenetrability of the atoms and the London dispersion forces, while the Coulombic potential describes the electrostatic interactions between charged atoms.

The connection between the force field and the dynamics of each atom i belonging to the molecular system is provided by the Newton equation of motion:

$$m_i \frac{d^2}{dt^2} \mathbf{r_i} = m \mathbf{a_i} = \mathbf{F_i} = -\frac{\partial}{\partial \mathbf{r_i}} E(\mathbf{r}) \tag{1.4}$$

where $\mathbf{F_i}$ is the force acting on atom i and $\mathbf{a_i}$ and m_i are the acceleration and the mass of atom i, respectively.

1.2 Solvation

Most macromolecules need to be solvated in an aqueous environment to express their function. As an active constituent in cell biology [9], water must be properly modeled to ensure the accuracy of MD simulations of biological processes. This is not an easy task, since up to now more than sixty anomalous thermodynamic and physical properties of water have been reported [10, 11]. Two approaches are commonly used to implement the water description in MD investigations: explicit solvent and implicit solvent representation. In the former, water molecules are explicitly described with the same atomistic resolution as the molecular system of interest. Several water models have been so far developed, which can be rigid or flexible, with or without treatment of polarization effects and with a variable number of interacting sites. In CHARMM, the TIP3P water model is usually employed, which is a three-site, non-polarizable model [12]. The explicit treatment of the water molecules is physically the most rigorous but has the disadvantage of dramatically increasing the computational cost. This hampers the study of long time scales processes like protein folding and aggregation, which require large scale structural rearrangements [13].

Implicit solvent models have been specifically developed in order to obtain an accurate but computationally cheap description of the aqueous environment [14, 15]. In this representation, the solute is immersed in a potential of mean force which models the mean influence of the water medium and depends only on the atom coordinates of the solute [13]. The use of implicit solvent models significantly reduces the system size and, in addition, lowers the viscosity arising from the friction of the water molecules, which results in speeding up the motion of the simulated molecules [16]. The hydration free energy of a molecule is usually decomposed in a polar term ΔG_{pol}, which takes into account the electrostatic interactions of water molecules with the charged atoms of the solute and the resulting screening of their Coulombic interactions, as well as a non-polar term ΔG_{npol}, which stems for the water cavitation term and for the van der Waals interaction between water and solute atoms. The scheme used to approximate these two terms characterize the different currently available implicit solvent models. Surface area models are based on the assumption that solvation energies can be decomposed into atomic contributions, which are governed by the local solute geometry [17, 18, 19, 20]. Continuum electrostatics approaches involve numerical solution of the Poisson-Boltzmann equation for the electrostatic potential $\phi(\mathbf{r})$:

$$\nabla \cdot [\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) \tag{1.5}$$

where $\epsilon(\mathbf{r})$ is the position dependent dielectric permittivity and $\rho(\mathbf{r})$ the charge distribution of the solute. This method allows accurate evaluation of the polar term of the hydration free energy and has been therefore implemented into some force fields [21], but its computational cost is still too high. Generalized Born (GB) models are related to the continuum electrostatics method but further simplifications enable the approximation of the Poisson-Boltzmann electrostatic solvation energy as an efficient pairwise potential. FACTS, an implicit solvent model used in some simulations reported in this thesis, is a GB model. In FACTS, a fully analytical evaluation of the volume and spatial symmetry of the solvent displaced from around a solute atom by its neighbors is performed to compute the hydration free energy [22].

1.3 Enhanced sampling techniques

MD simulations produce a set of configurations generated by iterative numerical solution of the Newton equation of motion (1.4). We say that the system "samples" different conformations during the simulation time. All relevant thermodynamic and structural observables will then be computed as an average over the sampled configurational space. This means that, if we are interested in studying a molecular process which involves large structural rearrangements taking place on the microsecond to second time scale, we have to perform a simulation of at least the same time length. Even more challenging is to obtain equilibrium measurements, since they require that the studied phenomenon happens reversibly during the simulation. Shaw and his coworkers showed that statistical convergence represents an issue even for a 200 μs explicit water MD simulation of the acid-unfolded state of the bovine acyl-coenzyme A binding protein, which is so far one of the longest MD run ever performed [23]. In the case of very small systems [24], the use of implicit solvent models can solve the problem of statistical convergence of the simulations, but in order to investigate highly complex processes like the folding of medium and large proteins, ligand association to intrinsically unstructured proteins [25] or protein self-assembly, techniques that speed up and improve the sampling have to be employed.

1.3.1 Coarse-grained models

One way of simplifying the complexity of a system is to eliminate atomistic details of the simulated biomolecules. In coarse-grained models, the atoms of a molecule are grouped into mesoscopic units or “beads”, which mutually interact through potentials of mean force [26]. For example, a methyl group can be represented with only one bead, if the investigator is not interested in the hydrogen dynamics. Polar hydrogens are, on the contrary, usually retained because they can participate in hydrogen bonds, as in the CHARMM 19 representation [27]. When a given atomistic description is simplified as in the previous examples, a “bottom-up” approach is applied. In another approach, a minimal set of interactions and degrees of freedom can be chosen upfront and the model parameters are then fitted to reproduce the essential features of the investigated process. This approach is called “top-down”. An example of a “top-down” model is the phenomenological coarse-grained model developed by Pellarin and Caflisch [28], which is extensively described in Chapter 2. In spite of the loss of high-resolution details, coarse-grained models based on both simplification methods have turned out to be particularly potent and useful to study nucleation-limited aggregation processes, because of the breadth in the time and space scales involved [29].

1.3.2 Replica exchange molecular dynamics simulations

Another commonly used enhanced sampling technique is the Replica Exchange Method (REM), which has been extended to MD simulations by Sugita and Okamoto [30]. The basic idea of this technique is to simulate in parallel different copies of the same system (*replicas*) at different temperature values. Each replica evolves independently but states belonging to neighboring temperatures are periodically exchanged with a probability which depends on the energy difference of the two states. Thus, the system performs a random walk in the temperature space, which is equivalent to the energy space. In this way, low temperatures simulation segments explore in details local minima, while the high temperatures portions of the trajectories help to cross kinetic barriers, so that the system does not get stuck in local traps on the free-energy surface. Replica exchange molecular dynamics (REMD) simulations have been successfully applied to study protein folding [31] and peptide aggregation [32] and have proven to be more efficient than constant temperature MD (CTMD) simulations in sampling the conformational space of a molecular system. However, the kinetic information is available only for short intervals of the trajectories, i.e., only between two temperature exchanges.

1.4 Multiscale computational approaches tailored for different investigation targets

Since the first MD simulation study [2], numerous and various techniques have been developed to to study increasingly complex systems with the basic original idea of numerically solving the equation of motion for the constituents of the studied molecular system. Each method relies on certain approximations, whose validity depends intrinsically on the target of study. In this thesis, different computational approaches of different resolution and complexity levels were tailored to different biological problems. In Chapter 2, an in-house developed coarse-grained phenomenological (CGF) model to study amyloid aggregation is extensively described and its main applications are reviewed. Chapter 3 deals with the use of the CGF model to investigate the effects of macromolecular crowding on the kinetics of self-assembly of a model peptide. This study was motivated by the observation that amyloid aggregation usually takes place in the cellular milieu, which is a crowded environment [33, 34]. Nonetheless, most of the experimental and computational investigations on protein self-assembly are usually performed in ideal homogeneous solutions. Molecular dynamics CTMD simulations with the FACTS generalized Born implicit solvent model were performed to obtain atomistic details of the structure of the early aggregates of phenylalanine, whose fibrillar assemblies have been hypothesized to be responsible for the neuronal damage observed in individuals affected by Phenylketonuria. The results of the framework experimental study and of our computational contribution to it are shown in Chapter 4. The influence of carnosine (β-alanyl-L-histidine), a dipeptide naturally occurring in the brain, on the conformational features of the full-length amyloid peptide Aβ_{1-42} has been studied with REMD simulations performed with the FACTS implicit solvent model. Moreover, carnosine's potential role as aggregation inhibitor has been investigated with CTMD simulations of aggregating Aβ segment peptides and carnosine, using FACTS to represent the aqueous environment. Chapter 5 presents our computational results, as well as the data obtained by our collaborators with different experimental techniques. Explicit water CTMD simulations have been used to study the binding site flexibility of 20 human bromodomains, protein modules which specifically recognize acetylated lysine side chains in histones and are therefore supposed to play a role in transcriptional regulation [35, 36]. Furthermore, we employed the same computational approach to characterize the mechanism and kinetics of acetyl-lsyine binding to the the human Transcription initiation factor TFIID subunit 1 (abbreviated TAF1(2)). The results on acetyl-lysine binding to TAF1(2) and binding site plasticity of bromodomains are shown in Chapter 6 and Chapter 7, respectively.

Bibliography

[1] Allen, M. and Tildesley, D. Computer Simulation of Liquids. Oxford University Press, 1989.

[2] Alder, B. J. and Wainwright, T. E. Phase transition for a hard sphere system. The Journal of Chemical Physics 27(5):1208–1209, 1957.

[3] van Gunsteren, W. F. and Berendsen, H. J. C. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. Angewandte Chemie International Edition in English 29(9):992–1023, 1990.

[4] Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. Journal of Computational Chemistry 4(2):187–217, 1983.

[5] Brooks, B. R., Brooks, C. L., MacKerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. CHARMM: the biomolecular simulation program. Journal of Computational Chemistry 30(10):1545–1614, 2009.

[6] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. A second generation

force field for the simulation of proteins, nucleic acids, and organic molecules. Journal of the American Chemical Society 117(19):5179–5197, 1995.

[7] Jorgensen, W. L. and Tirado-Rives, J. The OPLS (optimized potentials for liquid simulations) potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. Journal of the American Chemical Society 110(6):1657–1666, 1988.

[8] Oostenbrink, C., Villa, A., Mark, A. E., and Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The gromos force-field parameter sets 53a5 and 53a6. Journal of Computational Chemistry 25(13):1656–1676, 2004.

[9] Ball, P. Water as an active constituent in cell biology. Chemical Reviews 108(1):74–108, 2008.

[10] Chaplin, M. http://www.lsbu.ac.uk/water/anmlies.html.

[11] Debenedetti, P. G. Supercooled and glassy water. Journal of Physics: Condensed Matter 15(45):R1669, 2003.

[12] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics 79(2):926–935, 1983.

[13] Marchand, F. and Caflisch, A. Fast Analytical Continuum Treatments of Solvation, 209–232. Wiley-VCH Verlag GmbH & Co. KGaA, 2010.

[14] Feig, M. and Brooks, C. L. Recent advances in the development and application of implicit solvent models in biomolecule simulations. Current Opinion in Structural Biology 14(2):217–224, 2004.

[15] Roux, B. and Simonson, T. Implicit solvent models. Biophysical Chemistry 78(1–2):1–20, 1999.

[16] Zagrovic, B. and Pande, V. Solvent viscosity dependence of the folding rate of a small protein: Distributed computing study. Journal of Computational Chemistry 24(12):1432–1436, 2003.

[17] Eisenberg, D. and McLachlan, A. D. Solvation energy in protein folding and binding. Nature 319:199–203, 1986.

[18] Ooi, T., Oobatake, M., Némethy, G., and Scheraga, H. A. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. Proceedings of the National Academy of Sciences 84(10):3086–3090, 1987.

[19] Fraternali, F. and van Gunsteren, W. An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. Journal of Molecular Biology 256(5):939–948, 1996.

[20] Ferrara, P., Apostolakis, J., and Caflisch, A. Evaluation of a fast implicit solvent model for molecular dynamics simulations. Proteins: Structure, Function, and Bioinformatics 46(1):24–33, 2002.

[21] Nina, M., Im, W., and Roux, B. Optimized atomic radii for protein continuum electrostatics solvation forces. Biophysical Chemistry 78(1–2):89–96, 1999.

[22] Haberthür, U. and Caflisch, A. FACTS: Fast analytical continuum treatment of solvation. Journal of Computational Chemistry 29(5):701–715, 2008.

[23] Lindorff-Larsen, K., Trbovic, N., Maragakis, P., Piana, S., and Shaw, D. E. Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. Journal of the American Chemical Society 134(8):3787–3791, 2012.

[24] Krivov, S. V., Muff, S., Caflisch, A., and Karplus, M. One-dimensional barrier-preserving free-energy projections of a β-sheet miniprotein: New insights into the folding process. Journal of Physical Chemistry B 112(29):8701–8714, 2008.

[25] Wright, P. E. and Dyson, H. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. Journal of Molecular Biology 293(2):321–331, 1999.

[26] Müller, M., Katsov, K., and Schick, M. Biological and synthetic membranes: What can be learned from a coarse-grained description? Physics Reports 434(5–6):113–176, 2006.

[27] Reiher, W. Theoretical studies of hydrogen bonding. Ph.D. Thesis, Department of Chemistry, Harvard University, Cambridge, MA, USA, 1985.

[28] Pellarin, R. and Caflisch, A. Interpreting the aggregation kinetics of amyloid peptides. Journal of Molecular Biology 360(4):882–892, 2006.

[29] Wu, C. and Shea, J.-E. Coarse-grained models for protein aggregation. Current Opinion in Structural Biology 21(2):209–220, 2011.

[30] Sugita, Y. and Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. Chemical Physics Letters 314(1–2):141–151, 1999.

[31] Rao, F. and Caflisch, A. Replica exchange molecular dynamics simulations of reversible folding. Journal of Chemical Physics 119(7):4035–4042, 2003.

[32] Cecchini, M., Rao, F., Seeber, M., and Caflisch, A. Replica exchange molecular dynamics simulations of amyloid peptide aggregation. Journal of Chemical Physics 121(21):10748–10756, 2004.

[33] Fulton, A. B. How crowded is the cytoplasm? Cell 30(2):345–347, 1982.

[34] Zimmerman, S. B. and Trach, S. O. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of escherichia coli. Journal of Molecular Biology 222(3):599–620, 1991.

[35] Zeng, L. and Zhou, M.-M. Bromodomain: an acetyl-lysine binding domain. FEBS Letters 513(1):124–128, 2002.

[36] Filippakopoulos, P. and Knapp, S. The bromodomain interaction module. FEBS Letters 586(17):2692 – 2704, 2012.

Chapter 2

Mechanisms and kinetics of amyloid aggregation investigated by a phenomenological coarse-grained model

Magno, A., Caflisch, A. and Pellarin, R. Chapter of the book: *"Computational Modeling of Biological Systems: From Molecules to Pathways"*, **2012**, *191-214*, Springer

Mechanisms and kinetics of amyloid aggregation investigated by a phenomenological coarse-grained model

Andrea Magno, Riccardo Pellarin, and Amedeo Caflisch

[1]*Department of Biochemistry, University of Zürich,*

Winterthurerstrasse 190, CH-8057, Zürich, Switzerland[*]

I. INTRODUCTION

Amyloid fibrils are ordered polypeptide aggregates that have been implicated in several neurodegenerative pathologies, such as Alzheimer's, Parkinson's, Huntington's and prion diseases[1,2], and, more recently, also in biological functionalities[3–5]. These findings have paved the way for a wide range of experimental and computational studies aimed at understanding the details of the fibril formation mechanism. Computer simulations using low-resolution models, which employ a simplified representation of protein geometry and energetics, have provided insights into the basic physical principles underlying protein aggregation in general[6–8] and ordered amyloid aggregation[9–15]. For example, Dokholyan and coworkers have used the Discrete Molecular Dynamics method[16,17] to shed light on the mechanisms of protein oligomerization[18] and the conformational changes that take place in proteins before the aggregation onset[19,20]. One challenging observation, which is difficult to observe by computer simulations, is the wide range of aggregation scenarios emerging from a variety of biophysical measurements[21,22]. Atomistic models have been employed to study the conformational space of amyloidogenic polypeptides in the monomeric state[23–25], the very initial steps of amyloid formation[26–32], and the structural stability of fibril models[33–35]. However, all-atom simulations of the kinetics of fibril formation are beyond what can be done with modern computers.

To overcome such computational limitations, simplified models have been developed and used to investigate the kinetics and pathways of oligomerization and fibril formation at different levels of resolution[36]. In this chapter we first review briefly the simplified models of aggregation. We then present our coarse-grained phenomenological (CGF) model of an amphipathic peptide[37], and its use for studying kinetics and thermodynamics, both in bulk conditions and in presence of other simplified (macro)molecules.

II. COARSE-GRAINED MODELS

In coarse-grained models, the complexity of a system (and therefore the computational cost) is reduced by grouping atoms into larger units or "beads", whose mutual interactions are usually approximated by a potential of mean force[38]. Several coarse-grained models of different resolutions have been developed to study aggregation (see Fig.1). Zhang and Muthukumar[39] have created a cuboid model able of reproducing the features of a nucleation-limited aggregation process. With their so-called "tube" model, Auer and coworkers[40] have shed light upon the conversion of a disordered aggregate into an aggregating nucleus. Higher resolution models like the one developed by Thirumalai and collaborators[41] or the PRIME model of Nguyen and Hall[42] have also showed disordered aggregates in the early steps of aggregation. Several aggregation scenarios have been described with the three-bead model of Shea and coworkers[43]. In their model, the variation of a parameter related to the dihedral flexibility is able to reproduce different aggregation kinetics and metastable intermediates (amorphous and β-barrel like), which is in part similar to the CGF model[37]. The main difference between the Shea and CGF models is that the former is based on a coarse-graining from an atomistic description (i.e., "bottom-up" development), whereas the CGF model is purely phenomenological ("top-down") as explained in the next section.

III. THE CGF (COARSE-GRAINED PHENOMENOLOGICAL) MODEL

The coarse-grained model of an amphipathic peptide developed for studying aggregation kinetics and thermodynamics is a compromise between mesoscopic detail and computational efficiency. It must be stressed that this simplified model does not represent a particular protein sequence, i.e., it has not been generated by grouping into larger beads the atomic structure of a given (poly)peptide. It rather was designed from scratch for emulating the main experimental findings on fibril formation kinetics.

The peptide is approximated by 10 spherical beads, four of which represent the "backbone" (small beads) and six the "side chain" (large beads) (Fig. 2, top). The "backbone" beads carry partial charges of $\pm 0.4e$, thereby generating two dipoles; this part of the monomer is designed to interact specifically by intermolecular dipole-dipole interactions. The large beads interact only by van der Waals forces. The non-bonding interaction cutoff is set equal to 20 Å. The monomer displays an amphipathic moment, since eight of the ten beads have less favorable van der Waals interactions than the remaining two beads (black spheres in Fig. 2, top). The amphipathicity of the "molecule" allows the formation of amorphous aggregate, such as micellar oligomers, and the assembly of fibrils. In both of these types of aggregates, the hydrophobic spheres are buried and the hydrophilic spheres are exposed. The micellar oligomers are spherical and fluid-like, while the fibrils are ordered and rigid (see below).

The monomer can change its conformation by rotating around the internal dihedral defined by the small beads (Fig. 2, bottom). Using a one-dimensional spline function[44], a dihedral potential was designed with only two minima separated by a barrier (see Fig. 3). The only parameter that rules the relative populations of the amyloid-prone and amyloid-protected states is the energy difference $dE = E_\pi - E_\beta$ between the conformation with perpendicular dipoles (E_π), which prevents ordered aggregation, and the conformation with parallel dipoles, which is prone to form fibrils (E_β). The use of a single parameter to model a complex process was inspired by the work of Zhou and Karplus, who have analyzed the folding kinetics of a model protein by varying a single parameter and shown that it is possible to recover several folding scenarios[45].

IV. AGGREGATION OF THE CGF PEPTIDE MODEL IN BULK SOLUTION

Unless specified explicitly, simulations are started from 125 monodispersed monomers of the CGF peptide in a cubic box with a size of 290 Å, corresponding to a concentration of 8.5 mM. After minimization and equilibration, simulations are performed with Langevin dynamics at 310 K with a very small friction coefficient of 0.01 ps^{-1} using CHARMM[46].

A. Aggregation kinetics and pathways

The range of aggregation kinetics of the CGF model is shown in Fig. 4, where the normalized degree of polymerization as a function of time is plotted for different values of the amyloidogenic propensity dE. The extent of aggregation is controlled by counting the number of polar contacts: a polar contact is formed whenever two dipoles of different monomers are closer than 5 Å. Three different kinetic phases are visible: lag, elongation, and final monomer-fibril equilibrium. The variable length of the lag phase and the higher heterogeneity at longer lag times are indicative of a stochastic nucleation[48]. Fibril formation is much slower for the β-unstable models (dE = -2.5 and -2.25 kcal/mol) than the β-stable models (dE = -1.0 and 0.0 kcal/mol). Both the lag phase and the elongation kinetics are affected by the single free parameter dE of the CGF model. Interestingly, the kinetics of aggregation are qualitatively consistent with the experimental data on single-point mutants of Aβ_{40} (Ref.[49]), which have shown that the β-sheet propensity and hydrophobicity affect the features of the aggregation process. This comparison shows that although the CGF model does not represent any particular polypeptide sequence, variations of the single parameter dE emulate the behavior observed for (slightly) different amyloidogenic sequences. Moreover, the anticorrelation between the length of the lag phase and the rapidity of the fibril elongation has also been observed experimentally on several samples prepared from amyloidogenic (poly)peptide sequences[47].

The distribution function $p(N)$ of the oligomer size N evaluated at the lag phase or at the final equilibrium is depicted in Fig. 5. The monomer peak ranges from $N = 1$ to $N = 7$, the micellar peak from $N = 8$ to $N = 60$, and the fibrillar peak from $N = 61$ to $N = 125$. The micellar peak is present for the dE = -2.5 kcal/mol model at the lag phase, but disappears at the final equilibrium, where the fibril and the monomers are the only co-existing species. For the β-stable potential $dE = 0.0$ kcal/mol, the micellar peak is not observed at any concentration value. Indeed, a comparison of the lag times with the times of micelle formation (Fig. 6) shows that the fibril formation kinetics of the β-unstable and β-stable models are, respectively, slower and faster than micelle formation. In fact, micelles are intermediates consisting mainly of monomers in the π state, whereas the polymerization of β-stable monomers directly yields fibrils.

This observation is confirmed also by the analysis of aggregation pathways[50]. A total of 100 Langevin dynamics simulations for different values of dE were clustered according to three progress variables: the size of the largest aggregate N_{la}, the number of monomers in the β-state within the largest aggregate N_{la}^{β}, and the number of protofilaments in the largest aggregate N_{la}^{pf}, where a protofilament is defined as a file of monomers with intermolecular dipolar interactions parallel with its axis. The aggregation state network (Fig. 7) is a graph in which the states and direct transitions observed during the Langevin dynamics simulations are displayed as nodes and links, respectively. Furthermore, the size of each node reflects the statistical weight of the corresponding state. Micellar oligomers (white nodes, $N_{la} \sim 20$, N_{la}^{pf}=0) and fibrils (red nodes, $N_{la} \sim 100$, N_{la}^{pf}=4) are the most populated states during the lag phase and the final equilibrium, respectively. Strikingly, a greater variety of aggregation mechanisms emerges for the poorly amyloidogenic CGF peptide model (see Fig. 7, bottom) than the highly amyloidogenic CGF peptide model (see Fig. 7, top). Indeed, the former shows the presence of intermediates, i.e., protofibrils consisting of only two (green nodes) or three (blue nodes) protofilaments. According to this analysis, it is reasonable to expect that a mutation that decreases the β-aggregation tendency could result in a greater variety of prefibrillar aggregates, as in the case of the Arctic mutant (E22G) of the Alzheimer's Aβ peptide and the A30P mutant for α-synuclein, for which a more pronounced *in vitro* formation of oligomers and protofibrils was observed[51,52].

B. Mechanism of nucleation

The nucleation properties of the CGF model are investigated by evaluating the probability of fibril formation for β-subdomains, i.e., the clusters of interacting β-monomers.

The nucleus, defined as the oligomer containing a β-subdomain with a 50% probability to form a fibril, shows an increasing size upon destabilization of the β-state. Significantly different nucleation mechanisms are observed upon variation of the amyloidogenicity parameter dE (Fig. 8). For high values of the amyloidogenic propensity ($-2.0 \leq dE \leq 0.0$ kcal/mol), the nucleus size is sub-micellar, and nucleation is simply the aggregation of monomers in the β-state. On the contrary, for poorly amyloidogenic peptides, nucleus formation requires either spatial proximity of several monomers in the β-state ($dE =$ -2.25 kcal/mol) within a micelle or collision of two peptide micelles with merging of their β-subdomains ($dE =$ -2.5 kcal/mol). The variety of aggregation scenarios is also observed experimentally. An unstructured peptide with a marginally stable β-prone state like $A\beta_{40}$ (Ref.[53,54]) visits oligomeric intermediates in the lag phase, and has a very weak dependence of the elongation rate on concentration due to the monomer-micelle equilibrium. This mechanism corresponds to the nucleated conformational conversion proposed by Serio *et al.*[55]. On the other hand, a functional and non-pathological amyloid in mammalians[56] lacks on-pathway intermediates and corresponds to the highly amyloidogenic CGF peptide model. Once more, by varying the only free parameter dE of the CGF model, it is possible to describe the aggregation properties of a wide and diverse range of (poly)peptide sequences.

C. Concentration effects

The dE parameter of the CGF model has a strong influence on the concentration dependence of the fibril formation kinetics. In agreement with the above-mentioned mechanism of nucleation, CGF peptides poorly prone to aggregation nucleate only at concentration values larger than the critical concentration of peptide micelle formation, whereas CGF peptides with a high value of amyloidogenicity nucleate even at lower concentrations (Fig. 9, left). Furthermore, the dependence of the elongation rate on the concentration is only marginal at low amyloidogenic tendency (Fig. 9 (b)). The reduced concentration dependence originates from competitive polymerizations, i.e., the elongation of the fibril and the presence of micellar oligomers. This observation is a consequence of the monomer-micelle equilibrium of the CGF peptide model, which maintains a nearly constant concentration of isolated monomers[57].

D. Amyloid fibril polymorphism

Experiments based on electron and atomic force microscopy as well as solid-state NMR spectroscopy revealed that changing the samples conditions, such as the pH[59] or the cosolvent concentration[60], or introducing a mechanical perturbation[61,62] results in different amyloid fibril morphologies. Furthermore, even within the same sample, a number of coexisting morphologies can be detected[59,63]. Recently, it was observed that the CGF peptide model is able to generate fibrils with distinct morphologies[64]. Interestingly, the populations of the different morphologies are strongly and nontrivially influenced by the amyloidogenic propensity dE, and two main mechanisms for fibril morphogenesis emerge. When the CGF peptide is highly prone to aggregate (dE = -1.5, -2.0 kcal/mol), the morphogenesis is under thermodynamic control, meaning that the morphology with the highest stability will emerge with the highest probability. In contrast, when the CGF peptide has a low amyloidogenic tendency (dE = -2.25, -2.5 kcal/mol), the fibril morphogenesis is under kinetic control. The morphologies that nucleate more readily are not necessarily the most stable ones, but those whose precursors are kinetically more accessible, as revealed by the free energy profiles of the fibrillation[64]. For the low amyloidogenic scenario, the process of morphology differentiation can be represented by a branched tree (Fig. 10). During the lag phase, the micellar oligomers are in equilibrium with the dispersed monomers. The early morphology differentiation occurs at the nucleation step, where the formation of the protofibrillar intermediates is regulated by the structural bifurcation of the nucleus. The 2PP and 3PP1 intermediates are competent to 4PF1 fibrils, while the 3PP2 intermediate is competent to 4PF2(+,-) fibrils. Alternatively, the presence of 3PP2+ and 3PP2- intermediates that are directly competent to 4PF2+ and 4PF2- fibrils, respectively, has been observed, although these pathways were not quantitatively analyzed. Finally, the pathway of formation of 4PF3 fibrils was not investigated in detail, due to the small number of nucleation events of this morphology. The multiple-pathways process observed here has a close similarity with the scenario described by Goldsbury *et al.*[65], where two different morphologies of Aβ have distinct maturation pathways, either with or without the presence of metastable protofibrils.

V. AGGREGATION IN THE PRESENCE OF LIPID VESICLES AND INERT CROWDERS

Amyloid aggregation *in vivo* does not occur in bulk solution. Rather, it takes place in the extracellular space, whose composition includes metabolites and proteins, or within the cell, which is usually densely occupied by (macro)molecules like proteins, nucleic acids, and polysaccharides, as well as macromolecular assemblies and organelles[66]. Several research groups have investigated the interactions between lipid vesicles and amyloid aggregates[67–69], whose accumulation on the surface of lipid bilayers was observed to cause membrane damage. Aggregation has also been studied in crowded media[70–72], where the thermodynamics and kinetics of aggregation are expected to sensibly change.

A. Effect of lipid bilayers on CGF peptide aggregation

A three-bead model of a lipid molecule has been developed to study the CGF peptide model aggregation kinetics in the presence of a lipid vesicle[73,74]. Several independent Langevin simulations at 310 K have been performed for four values of dE with 125 peptides initially monodispersed in a cubic box of length 290 Å and a preequilibrated unilamellar bilayer vesicle made up of 1000 lipids. Depending on the lipid/peptide van der Waals coupling parameter λ, between 50% and 80% of the CGF peptides are located on the lipid vesicle surface after the initial equilibration phase, i.e., before fibril formation (Table 1).

The effect of lipid bilayers on aggregation kinetics for different values of amyloidogenicity is reported in Table 2. Highly amyloidogenic peptides fibrillate more rapidly in the presence of lipid vesicles than in their absence, while the opposite is observed for peptides of low amyloidogenicity. The faster aggregation kinetics of highly amyloidogenic peptides is a consequence of their higher effective concentration on the lipid bilayer relative to the bulk. In contrast, despite the same increase of peptide concentration on the vesicle surface, fibrillation of peptides with low amyloidogenic propensity is slower in the presence of lipid vesicles. As mentioned in section III, peptides with low amyloidogenic potential can fibrillate only after aggregating into spherical oligomeric intermediates with hydrophobic interior and hydrophilic surface. In the simulations with lipids, such oligomeric intermediates form in the bulk but not on the vesicle. Fibrillation of low amyloidogenic peptides therefore takes place in the bulk and is slower than in the absence of a vesicle due to the lower effective concentration of peptides in the solvent. These simulation results are consistent with and explain the apparently contradictory experimental observations on faster aggregation of the Aβ(Ref.[68]) or α-synuclein[75] peptides in the presence of lipid surfaces and slower aggregation of insulin (which has lower amyloidogenicity)[76], and have been confirmed by recent studies on the aggregation properties of human islet amyloid polypeptide hIAPP$_{1-19}$ in presence of lipid vesicles[77].

To investigate the influence of the CGF peptides on the lipid bilayer, the simulations were initiated with 20 spherical probes inside the vesicle. It was observed that leakage from the lipid vesicle is enhanced during fibril formation but not by the mature fibril[73]. More precisely, a comparison between the fibrillation and probe release rates (Fig. 11, left) revealed that probe release is fastest during fibril growth, whereas the kinetics of probe release in the presence of mature fibrils is as slow as in the absence of peptides, indicating that mature fibrils do not damage the integrity of the vesicle. Rather, the ongoing process of aggregation on the vesicle results in bilayer surface defects. This observation explains why for some amyloidogenic peptides there exist mutants that form fibrils more rapidly and are more toxic than the wild-type peptides, even though their fibrils are not toxic[78]. Moreover, these computational results are in agreement with the experiments performed by Engel *et al.* on membrane damage caused by human islet amyloid polypeptide fibril growth[69] (Fig. 11 right).

It has also been hypothesized that formation of toxic oligomers that induce membrane leakage could be the result of a backward production of oligomers from the mature fibril[79]. Interestingly, by modulating the attraction between the CGF peptides and the membrane, fibril disaggregation into soluble backwards oligomers has been observed[74]. The disaggregation process is driven by entropy and results in soluble protofibrillar oligomers. The protofibrillar oligomers are larger, more ordered, and more stable than those observed during the aggregation process and, importantly, are not detected in disaggregation simulations carried out in bulk solution, i.e., in the absence of lipid vesicles.

B. Effect of surfactants on CGF peptide aggregation

Surfactant molecules have been modeled using a similar three-bead model as that used for lipids. The surfactant model differs from lipid models used previously in two parameters (Table 1): the minimum of the van der Waals energy of the two hydrophobic beads is less favorable, and the radius of the hydrophilic bead is larger to enable the formation of amorphous aggregates. Using these parameters, the surfactant solution is not dominated by a micellar phase. Rather, the surfactants are organized either as dispersed monomers or disordered aggregates[80].

In the absence of surfactants, all peptide models form fibrils within 1 μs without any discernible lag phase (Fig. 12, dotted lines). At a surfactant:peptide ratio of 8:1, the fibril formation kinetics of peptides with dE = -1.5 kcal/mol are almost unaffected, whereas already at a ratio of 4:1 the ordered self-assembly of peptides with $dE <$ -2.0 kcal/mol is significantly slower (Fig. 12, solid lines) mainly because of a longer lag phase. Moreover, for low-amyloidogenic peptides ($dE \leq$ -2.0 kcal/mol) no fibrillation is observed within the simulation length of 2 μs at a surfactant:peptide ratio of 8:1, but instead oligomers of $\sim$ 40 peptides form. These simulation results show that at a 4-fold molar excess of surfactant, the inhibition of fibrillation already depends strongly on the amyloidogenicity of the CGF peptide model.

C. Macromolecular crowding effect on CGF peptide aggregation

Simulations with the CGF peptide model together with softly repulsive spheres have been carried out to assess the influence on the aggregation kinetics of excluded volume and hindered peptide diffusion due to macromolecular crowding [81]. As in the case of lipid-bilayer vesicles, the net effect of macromolecular crowding crucially depends on the amyloidogenicity tendency of the CGF peptide. For peptides with low aggregation propensity, the self-association process is transition-state limited, where the kinetic bottleneck is the formation of the fibril nucleus. In this case, since the oligomers, including the nucleus, are thermodynamically favored (with respect to the isolated monomers) by the excluded volume effect, macromolecular crowding accelerates peptide assembly and has an effect analogous to that of an increase in peptide concentration (Fig. 13, left). This trend is analogous to that observed experimentally by Munishkina *et al.*, who have studied the effect of increasing the PEG concentration on the α-synuclein aggregation process[71].

On the other hand, when the aggregation mechanism is fast and proceeds directly from monomers to fibril, the process is diffusion limited, and the thermodynamic stabilization of oligomers is less important than the reduction in peptide mobility. In this case, the bottleneck is not the formation of the nucleus; the rate-limiting step for peptides that show a direct aggregation mechanism is the elongation of the fibril. Therefore, in this case macromolecular crowding is much less efficient in accelerating the self-association of peptides than an equivalent increase in peptide concentration, since the peptides diffusion is hindered by the crowders (Fig.13, right).

VI. CONCLUSION

Atomistic simulations of aggregation are limited by short time scale, while experimental approaches to amyloid fibril formation have insufficient spatial resolution. Coarse-grained models of polypeptide aggregation sacrifice atomistic detail to reach timescales that allow the comparison with and interpretation of experimental data. The models presented in this chapter have shed light upon amyloid aggregation kinetics and mechanisms, which is helpful to formulate a unified picture of the available experimental data.

The CGF model has only one tunable parameter, the difference dE between the energy of the amyloid-competent and the amyloid-protected state of the monomer[37]. Variations of this parameter reproduce several aggregation scenarios, both under homogeneous and heterogeneous conditions. It is important to highlight that the CGF model does not mimic any particular amyloid (poly)peptide sequence. However, the different aggregation kinetics obtained with this model can be directly compared with experiments carried out with specific proteins. In Table 3 are reviewed the principal characteristics of the aggregation process for both the high and low amyloidogenic tendency, and in both cases several examples of real amyloid-forming (poly)peptide sequences are listed. It is important to note that (coarse-grained) simulations, e.g., those with the CGF[37] and Shea[43] models, allow for the emulation of conditions and/or phenomena that are not accessible by (standard) experiments. As an example, the possibility to change solely the intrinsic conformational landscape of a monomer without affecting the inter-monomer interactions is an advantage of the (coarse-grained) simulation methods with respect to conventional experimental techniques such as mutagenesis and solvent-induced conformational changes, by which it is not possible to decouple changes in intra from inter-molecular interactions.

In conclusion, a slight modification of the free energy profile of an extremely simplified model of an amphipathic peptide is sufficient to observe a wide range of different fibril formation mechanisms, providing a unifying description of the heterogeneity of the experimentally observed kinetics of amyloid fibril formation.

---

[*] Author to whom correspondence should be addressed: caflisch@bioc.uzh.ch

[1] C. M. Dobson, Protein Folding and Misfolding. *Nature*, **426**, 884-890 (2003).

[2] P. T. Lansbury, H. A. Lashuel, A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature*, **443**, 774-779 (2006).

[3] D. M. Fowler, A. V. Koulov, W. E. Balch, J. W. Kelly, Functional amyloid–from bacteria to

humans. *Trends Biochem. Sci.*, **32**, 217-224 (2007).

[4] S. K. Maji, M. H. Perrin, M. R. Sawaya, S. Jessberger, K. Vadodaria, R. A. Rissman, P. S. Singru, K. P. R. Nilsson, R. Simon, D. Schubert *et al.*, Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science*, **325**, 328-332 (2009).

[5] J. Greenwald, R. Riek, Biology of amyloid: structure, function, and regulation. *Structure*, **18**, 1244-1260 (2010).

[6] R. A. Broglia, G. Tiana, S. Pasquali, H. E. Roman, E. Vigezzi, Folding and aggregation of designed proteins. *Proc. Natl Acad. Sci.* USA, **95**, 12930-12933 (1998).

[7] P. Gupta, C. K. Hall, A. C. Voegler, Effect of denaturant and protein concentrations upon protein refolding and aggregation: a simple lattice model. *Protein Sci.*, **7**, 2642-2652 (1998).

[8] P. M. Harrison, H. S. Chan, S. B. Prusiner, F. E. Cohen, Thermodynamics of model prions and its implications for the problem of prion protein folding. *J. Mol. Biol.*, **286**, 593-606 (1999).

[9] B. Urbanc, L. Cruz, S. Yun, S. V. Buldyrev, G. Bitan, D. B. Teplow, H. E. Stanley, In silico study of amyloid beta-protein folding and oligomerization. *Proc. Natl Acad. Sci.* USA, **101**, 17345-17350 (2004).

[10] J. Sørensen, X. Periole, K. K. Skeby, S. -J. Marrink, B. Schiøtt, Protofibrillar Assembly Toward the Formation of Amyloid Fibrils. *J. Chem. Phys. Lett.*, **2**, 2385-2390 (2011).

[11] H. Jang, C. K. Hall, Y. Zhou, Assembly and kinetic folding pathways of a tetrameric beta-sheet complex: molecular dynamics simulations on simplified off-lattice protein models. *Biophys. J.*, **86** (1 Pt 1), 31-49 (2004).

[12] R. I. Dima, D. Thirumalai, Exploring protein aggregation and self-propagation using lattice models: phase diagram and kinetics. *Protein Sci.*, **11** (5), 1036-1049 (2002).

[13] E. Malolepsza, M. Boniecki, A. Kolinski, L. Piela, Theoretical model of prion propagation: a misfolded protein induces misfolding. *Proc. Natl Acad. Sci.* USA, **102**, 7835-7840 (2005).

[14] S. D. Khare, F. Ding, K. N. Gwanmesia, N. V. Dokholyan, Molecular origin of polyglutamine aggregation in neurodegenerative diseases. *PLoS Comput. Biol.*, **1**, 230-235 (2005).

[15] Y. Chen, N. V. Dokholyan, A single disulfide bond differentiates aggregation pathways of beta2-microglobulin. *J. Mol. Biol.*, **354**, 473-482 (2005).

[16] N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, E. I. Shakhnovich, Discrete molecular dynamics studies of the folding of a protein-like model. *Folding Des.*, **3**, 577587 (1998).

[17] F. Ding, S. V. Buldyrev, N. V. Dokholyan, Folding Trp-cage to NMR resolution native structure using a coarse-grained protein model. *Biophys. J.*, **88**, 147155 (2005).

[18] F. Ding, N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, E. I. Shakhnovic, Molecular

dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. *J. Mol. Biol.*, **324**, 851-857 (2002).

[19] F. Ding, J. M. Borreguero, S. V. Buldyrey, H. E. Stanley, N. V. Dokholyan, Mechanism for the alpha-helix to beta-hairpin transition. *Proteins*, **53**, 220-228 (2003).

[20] F. Ding, J. J. LaRocque, N. V. Dokholyan, Direct observation of protein folding, aggregation, and a prion-like conformational conversion. *J. Biol. Chem.*, **48**, 40235-40240 (2005).

[21] W. S. Gosal, I. J. Morten, E. W. Hewitt, D. A. Smith, N. H. Thomson, S. E. Radford, Competing pathways determine fibril morphology in the self-assembly of beta2-microglobulin into amyloid. *J. Mol. Biol.*, **351**, 850-864 (2005).

[22] G. Plakoutsi, F. Bemporad, M. Calamai, N. Taddei, C. M. Dobson, F. Chiti, Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. *J. Mol. Biol.*, **351**, 910-922 (2005).

[23] A. Vitalis, X. Wang, R. V. Pappu, Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories. *Biophys. J.*, **93**, 1923-1937 (2007).

[24] A. Vitalis, N. Lyle, R. V. Pappu, Thermodynamics of beta-sheet formation in polyglutamine. *Biophys. J.*, **97**, 303-311 (2009).

[25] A. Vitalis, A. Caflisch, Micelle-like architecture of the monomer ensemble of Alzheimer's amyloid- peptide in aqueous solution and its implications for Aβ aggregation. *J. Mol. Biol.*, **403**, 148-165 (2010).

[26] J. Gsponer, U. Haberthr, A. Caflisch, The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proc. Natl. Acad. Sci.* USA, **100**, 5154-5159 (2003).

[27] W. Hwang, S. Zhang, R. D. Kamm, M. Karplus, Kinetic control of dimer structure formation in amyloid fibrillogenesis. *Proc. Natl. Acad. Sci.* USA, **101**, 12916-12921 (2004).

[28] M. L. de la Paz, G. M. S. de Mori, L. Serrano, G. Colombo, Sequence dependence of amyloid fibril formation: insights from molecular dynamics simulations. *J. Mol. Biol.*, **349**, 583-596 (2005).

[29] M. Cecchini, R. Curcio, M. Pappalardo, R. Melki, A. Caflisch, A molecular dynamics approach to the structural characterization of amyloid aggregation. *J. Mol. Biol.*, **357**, 1306-1321 (2006).

[30] B. Strodel, C. S. Whittleston, D. J. Wales, Thermodynamics and kinetics of aggregation for the GNNQQNY peptide. *J. Am. Chem. Soc.*, **129**, 16005-16014 (2007).

[31] A. De Simone, L. Esposito, C. Pedone, L. Vitagliano, Insights into stability and toxicity of

amyloid-like oligomers by replica exchange molecular dynamics analyses. *Biophys. J.*, **95**, 1965-1973 (2008).

[32] G. Bellesia, J. -E. Shea, What determines the structure and stability of KFFE monomers, dimers, and protofibrils? *Biophys. J.*, **96**, 875-886 (2009).

[33] B. Ma, R. Nussinov, Stabilities and conformations of Alzheimer's beta -amyloid peptide oligomers (Abeta 16-22, Abeta 16-35, and Abeta 10-35): Sequence effects. *Proc. Natl. Acad. Sci.* USA, **99**, 14126-14131 (2002).

[34] N. -V. Buchete, R. Tycko, G. Hummer, Molecular dynamics simulations of Alzheimer's beta-amyloid protofilaments. *J. Mol. Biol.*, **353**, 804-821 (2005).

[35] C. Wu, M. T. Bowers, J. -E. Shea, Molecular structures of quiescently grown and brain-derived polymorphic fibrils of the Alzheimer amyloid abeta9-40 peptide: a comparison to agitated fibrils. *PLoS Comput. Biol.*, **6**, e1000693 (2010).

[36] C. Wu, J. -E. Shea, Coarse-grained models for protein aggregation. *COSB*, **21**, 209-220 (2011).

[37] R. Pellarin, A. Caflisch, Interpreting the aggregation kinetics of amyloid peptides. *J. Mol. Biol.*, **360**, 882-892 (2006).

[38] M. Müller, K. Katsov, M. Schick, Biological and synthetic membranes: What can be learned from a coarse-grained decription? *Physics Reports*, **434**, 113-176 (2006).

[39] J. Zhang, M. Muthukumar, Simulations of nucleation and elongation of amyloid fibrils. *J. Chem. Phys.*, **130**, 035102 (2009).

[40] S. Auer, C. M. Dobson, M. Vendruscolo, A. Maritan, Self-templated nucleation in peptide and protein aggregation. *PLos Comput. Biol.*, **4**, e1000222 (2008).

[41] M. S. Li, D. K. Klimov, J. E. Straub, D. Thirumalai, Probing the mechanisms of fibril formation using lattice models. *J. Chem. Phys*, **129**, 175101 (2008).

[42] H. D. Nguyen, C. K. Hall, Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc. Natl Acad. Sci.* USA, **101**, 16180-16185 (2004).

[43] G. Bellesia, J. -E. Shea, Self-assembly of beta-sheet forming peptides into chiral fibrillar aggregates. *J. Chem. Phys.*, **126**, 245104 (2007).

[44] A. D. J. MacKerell, M. Feig, C. L. Brooks, Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.*, **126**, 698-699 (2004).

[45] Y. Zhou. M. Karplus, Interpreting the folding kinetics of helical proteins. *Nature*, **401**, 400-403 (1999).

[46] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch *et al.*, CHARMM: the biomolecular simulation program *J*

Comput. Chem., **30**, 1545-1614 (2009).

[47] M. Fändrich, Absolute correlation between lag time and growth rate in the spontaneous formation of several amyloid-like aggregates and fibrils. *J. Mol. Biol.*, **365**, 1266-1270 (2007).

[48] P. Hortschansky, V. Schroeckh, T. Christopeit, G. Zandomeneghi, M. Fändrich, The aggregation kinetics of Alzheimer's beta-amyloid peptide is controlled by stochastic nucleation. *Protein Sci.*, **14**, 1753-1759 (2005).

[49] T. Christopeit, P. Hortschansky, V. Schroeckh, K. Guhrs, G. Zandomeneghi, .M Fändrich, Mutagenic analysis of the nucleation propensity of oxidized Alzheimer's beta-amyloid peptide. *Protein Sci.*, **14**, 2125-2131 (2005).

[50] R. Pellarin, E. Guarnera, A. Caflisch, Pathways and intermediates of amyloid fibril formation. *J. Mol. Biol.*, **374**, 917-924 (2007).

[51] C. Nilsberth, A. Westlind-Danielsson, C. B. Eckman, M. M. Condron, K. Axelman, C. Forsell *et al.*, The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. *Nature Neurosci.*, **4**, 887-893 (2001).

[52] K. A. Conway, S. J. Lee, J. C. Rochet, T. T. Ding, R. E. Williamson, P. T. Lansbury, Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy. *Proc. Natl Acad. Sci.* USA, **97**, 571-576 (2000).

[53] R. Sabate, J. Estelrich, Evidence of the existence of micelles in the fibrillogenesis of beta-amyloid peptide. *J. Phys. Chem. B*, **109**, 11027-11032 (2005).

[54] A. Lomakin, D. S. Chung, G. B. Benedek, D. A. Kirschner, D. B. Teplow, On the nucleation and growth of amyloid beta-protein fibrils: detection of nuclei and quantitation of rate constants. *Proc. Natl Acad. Sci.* USA, **93**, 1125-1129 (1996).

[55] T. R. Serio, A. G. Cashikar, A. S. Kowal, G. J. Sawicki, J. J. Moslehi, L. Serpell *et al.*, Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science*, **289**, 1317-1321 (2000).

[56] D. M. Fowler, A. V. Koulov, C. Alory-Jost, M. S. Marks, W. E. Balch, J. W. Kelly, Functional amyloid formation within mammalian tissue. *PLoS Biol.*, **4**, e6 (2006).

[57] A. Lomakin, D. B. Teplow, D. A. Kirschner, G. Benedek, Kinetic theory of fibrillogenesis of amyloid beta-protein. *Proc. Natl Acad. Sci.* USA, **94**, 7942-7947 (1997).

[58] L. Nielsen, R. Khurana, A. Coats, S. Frokjaer, J. Brange, S. Vyas, V. N. Uversky, A. L. Fink, Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. *Biochemistry*, **40**, 6036-6046 (2001).

[59] C. Wasmer, A. Soragni, R. Sabate, A. Lange, R. Riek, B. H. Meier, Infectious and nonin-

fectious amyloids of the HET-s(218-289) prion have different NMR spectra. *Angew. Chem.*, Int. Ed., **47**, 5839-5841 2008.

[60] W. Dzwolak, S. Grudzielanek, V. Smirnovas, R. Ravindra, C. Nicolini, R. Jansen, A. Loksztejn, S. Porowski, R. Winter, Ethanol-perturbed amyloidogenic self-assembly of insulin: looking for origins of amyloid strains. *Biochemistry*, **44**, 8948-8958 (2005).

[61] A. T. Petkova, R. D. Leapman, Z. Guo, W. -M. Yau, M. P. Mattson, R. Tycko, Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science*, **307**, 262-265 (2005).

[62] A. K. Paravastu, A. T. Petkova, R. Tycko, Polymorphic fibril formation by residues 10-40 of the Alzheimer's beta-amyloid peptide. *Biophys. J.*, **90**, 4618-4629 (2006).

[63] J. Meinhardt, C. Sachse, P. Hortschansky, N. Grigorieff, M. Fändrich, Abeta(1-40) fibril polymorphism implies diverse interaction patterns in amyloid fibrils.
J. Mol. Biol., **386**, 869-877 (2009).

[64] R. Pellarin, P. Schuetz, E. Guarnera, A. Caflisch, Amyloid fibril polymorphism is under kinetic control. *J. Am. Chem. Soc.*, **132**, 14960-14970 (2010).

[65] C. Goldsbury, P. Frey, V. Olivieri, U. Aebi, S. A. Müller, Multiple assembly pathways underlie amyloid-beta fibril polymorphisms. *J. Mol.Biol.*, **352**, 282-298 (2005).

[66] J. R. Ellis, Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597-604 (2001).

[67] D. Lopes, A. Meister, A. Gohlke, A. Hauser, A. Blume, R. Winter, Mechanism of islet amyloid polypeptide fibrillation at lipid interfaces studied by infrared reflection absorption spectroscopy. *Biophys. J.*, **93**, 3132-3141 (2007).

[68] E. Chi, C. Ege, A. Winans, J. Majewski, G. Wu, K. Kjaer, K. Lee, Lipid membrane templates the ordering and induces the fibrillogenesis of Alzheimer's disease amyloid-beta peptide. *Proteins*, **72**, 1-24 (2008).

[69] M. F. M. Engel, L. Khemtémourian, C. Kleijer, H. Meeldijk, J. Jacobs, A. Verkleij *et al.*, Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. *Proc. Natl Acad. Sci. USA*, **105**, 60336038 (2008).

[70] R. J. Ellis, A. P. Minton, Protein aggregation in crowded environments. *Biol. Chem.*, **387**, 485-497 (2006).

[71] L. A. Munishkina, E. M. Cooper, V. N. Uversky, A. L. Fink, The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit.*, **17**, 456-464 (2004).

[72] L. A. Munishkina, A. Ahmad, A. L. Fink, V. N. Uversky, Guiding protein aggregation with

macromolecular crowding. *Biochemistry* **47**, 8993-9006 (2008).

[73] R. Friedman, R. Pellarin, A. Caflisch, Amyloid aggregation on lipid bilayers and its impact on membrane permeability. *J. Mol. Biol.*, **387**, 407-415 (2009).

[74] R. Friedman, R. Pellarin, A. Caflisch, Soluble Protofibrils as Metastable Intermediates in Simulations of Amyloid Fibril Degradation Induced by Lipid Vesicles *J. Phys. Chem. Lett.*, **1**, 471-474 (2010).

[75] M. J. Volles, S. J. Lee, J. C. Rochet, M. D. Shtilerman, T. T. Ding, J. C. Kessler, P. T. Lansbury, Vesicle permeabilization by protofibrillar alpha-synuclein: implications for the pathogenesis and treatment of Parkinson's disease. *Biochemistry*, **40**, 7812-7819 (2001).

[76] J. Sharp, J. Forrest, R. Jones, Surface denaturation and amyloid fibril formation of insulin at model lipid-water interfaces. *Biochemistry*, **41**, 15810-15819 (2002).

[77] L. Khemtémourian, M. F. M. Engel, R. M. J. Liskamp, J. W. M. Höppener, J. A. Killian, The N-terminal fragment of human islet amyloid polypeptide is non-fibrillogenic in the presence of membranes and does not cause leakage of bilayers of physiologically relevant lipid composition. *Biochimica et Biophysica Acta-Biomembranes*, **1798**, 1805-1811 (2010).

[78] H. Lashuel, P. Lansbury, Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Q. Rev. Biophys.*, **39**, 167-201 (2006).

[79] I. C. Martins, I. Kuperstein, H. Wilkinson, E. Maes, M. Vanbrabant, W. Jonckheere, P. V. Gelder, D. Hartmann, R. D'Hooge, B. D. Strooper *et al.*, Lipids revert inert Abeta amyloid fibrils to neurotoxic protofibrils that affect learning in mice. *EMBO J*, **27**, 224-233 (2008).

[80] R. Friedman, A. Caflisch, Surfactant Effects on Amyloid Aggregation Kinetics. *J. Mol. Biol.*, **414**, 303-312 (2011).

[81] A. Magno, A. Caflisch, R. Pellarin, Crowding Effects on Amyloid Aggregation Kinetics. *J. Phys. Chem. Lett.*, **1**, 3027-3032 (2010).

Type of molecule	$R_{hydrophilic}$ [nm]	$R_{hydrophobic}$ [nm]	$\epsilon_{hydrophilic}$ [kcal/mol]	$\epsilon_{hydrophobic}$ [kcal/mol]	λ^a	Fraction of CGF peptides bound to lipid vesicles	Reference
lipid	0.31	0.3	-0.1	-1.265	0.87-0.90	50%	[73]
lipid	0.31	0.3	-0.1	-1.265	0.95	80%	[74]
surfactant	0.35	0.3	-0.1	-0.8	1		[80]

TABLE 1. Three-bead lipid and surfactants models used with CGF peptide model. ^{a}Scaling factor for the vdW interactions between lipids or surfactants and peptides. Different scaling is used to model different systems, i.e., surfactants ($\lambda = 1$), moderately attractive ($\lambda \leq 0.9$) and strongly attractive ($\lambda \approx 0.95$) lipid bilayers.

Scaling of peptide/lipid interactions	Amyloidogenicity	Number of runs with fibril formation	Lag time t_{50} [ns]	
			with membrane	without membrane
0.87	high	10/10	11 ± 1	19 ± 3
	interm.	29/29	89 ± 29	56 ± 15
	low	17/20	**958** ± 503	124 ± 28
	very low	0/10	> **2000**	318 ± 133
0.90	high	10/10	10 ± 1	19 ± 3
	interm.	30/30	69 ± 23	56 ± 15
	low	0/20	> **2000**	124 ± 28
	very low	0/20	> **2000**	318 ± 133

TABLE 2. Characteristic lag time of aggregation t_{50} for CGF peptide model for different amyloidogenic tendency in the presence or absence of lipid vesicles[73]. Values in boldface are significantly larger in the presence of the vesicles

high amyloidogenicity	low amyloidogenicity	Ref.
small nucleus	large nucleus	[37]
fast fibril formation	slow fibril formation	[37]
downhill	micellar intermediates	[37]
no intermediates	protofibrillar intermediates	[50]
single pathway	multiple pathways	[50]
strong concentration dependence	growth rate marginally dependent on concentration	[37]
polymorphism under thermodynamic control	polymorphism under kinetic control	[64]
can promote membrane leakage	does not promote membrane leakage	[73]
slightly accelerated by membranes	decelerated by membranes	[73]
marginally influenced by surfactants	decelerated by surfactants	[80]
not accelerated by macromolecular crowding	accelerated by macromolecular crowding	[81]
Phe-Phe, GNNQQNY, transthyretin, $A\beta_{42}$	$A\beta_{40}$, Sup35, prion protein, myoglobin	-

TABLE 3. Influence of amyloidogenic propensity on the aggregation kinetics and pathways of the CGF model[37]. The last line lists some examples but it must be stressed that amyloidogenic tendency strongly depends on external conditions, so that the same polypeptide sequence can show drastically different amyloidogenic tendency depending on pH, temperature, etc.

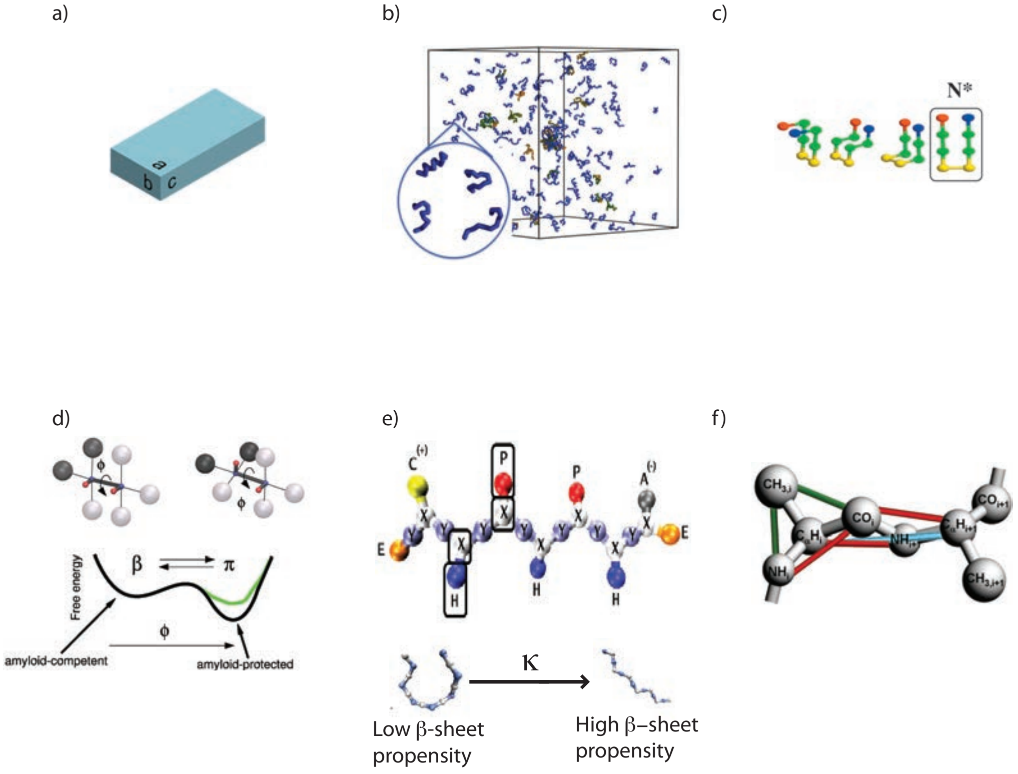


FIG. 1. Main coarse-grained models discussed in section II. a) Cuboid model[39]; b) tube model[40]; c) lattice model[41]; d) CGF model[37]; e) Shea model[43]; f) Hall model[42]. Reprinted from[36] with permission by Elsevier.

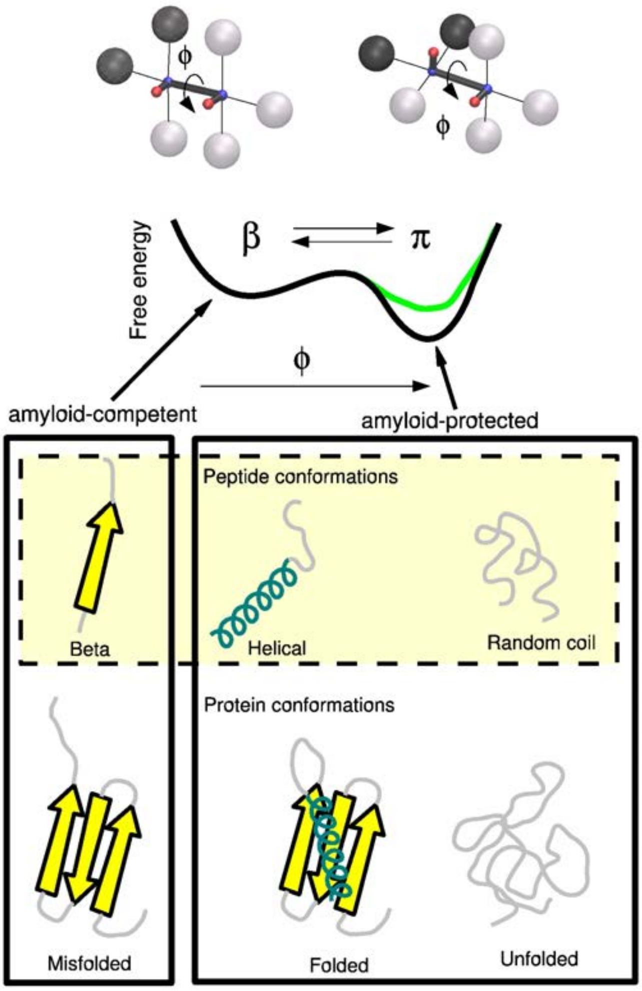


FIG. 2. The CGF model: sticks and beads representations of the monomer in the amyloid-competent state β and the amyloid-protected state π[37]. The large spheres are hydrophobic (black) and hydrophilic (gray), while the two dipoles are shown with small red and blue spheres. The size of the spheres does not represent the actual van der Waals radii, which are 2.5 Å for the black and gray spheres and 2.0 Å for the red and blue spheres. The β and π states of the monomer are shown on top of the two corresponding minima of the free energy, plotted as a function of the dihedral angle ϕ of the two dipoles. Reprinted from[50] with permission by Elsevier.

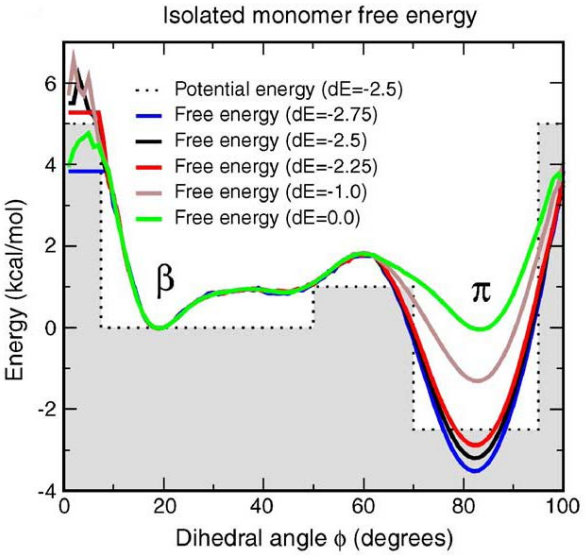
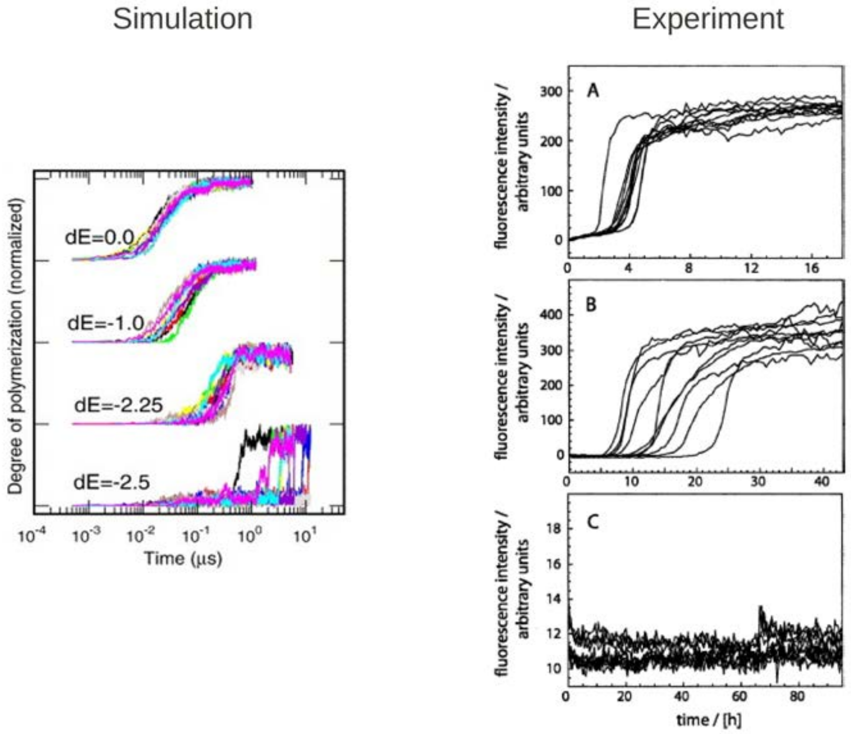


FIG. 3. The dotted line is the dihedral potential with a dE = -2.5 kcal/mol energy difference between amyloid-protected and amyloid-competent state. The five continuous lines represent the free energy profile of the isolated monomer for five different dihedral potentials. Since the peptide has only one degree of freedom, dE is close to the free-energy difference between the two aforementioned states. For instance, when dE = 0.0 kcal/mol, the π and β states are equally populated, whereas for dE = -1.5, -2.0, and -2.25 kcal/mol, the π state is about 15, 39, and 64 times more populated than the β state, respectively. Reprinted from[37] with permission by Elsevier.


Simulation

Experiment

FIG. 4. Influence of amyloidogenic tendency on aggregation kinetics. (Left): Time series of the fraction of ordered aggregation evaluated at four values of the amyloidogenic tendency, from very prone to form fibrillar aggregates (dE = 0.0 kcal/mol) to marginal propensity (dE = -2.5 kcal/mol). Ten independent simulations are shown for each dE value. (Right): Fluorescence intensity (degree of aggregation) of V18I (a), V18Q (b) and V18P (c) mutants of Aβ_{40} (Ref.[49]). Note that the ns-μs timescales in the CGF model simulations are much shorter than in the experiments (hours) because of the much higher concentration in the former (8.5 mM) than in the latter (120 μM). Reprinted from[37] (left) and[49] (right) with permission by Elsevier (left) and Jon Wiley & Sons (right).

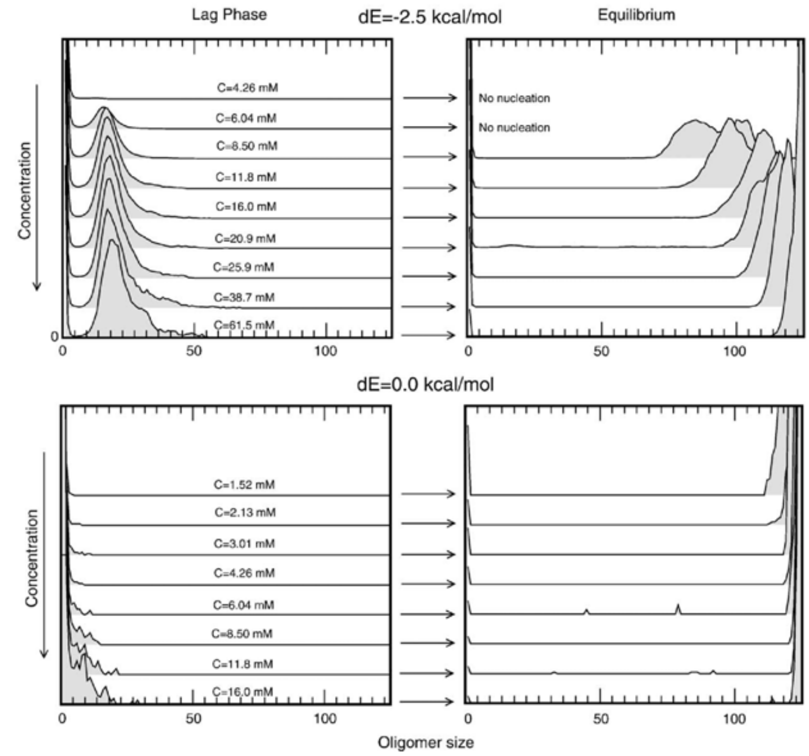


FIG. 5. Oligomer size histograms of the dE = -2.5 kcal/mol potential (top) and dE = 0.0 kcal/mol (bottom) calculated at the lag phase (left) and at the final equilibrium (right). The z-dimension represents the relative probability. Note that most of the results were obtained at a concentration C = 8.5 mM which is the lowest value at which fibril formation takes place within a reasonable simulation time for the dE = -2.5 kcal/mol model (10 μs in about 17 days on a single Xeon 5410 processor). Reprinted from[37] with permission by Elsevier.

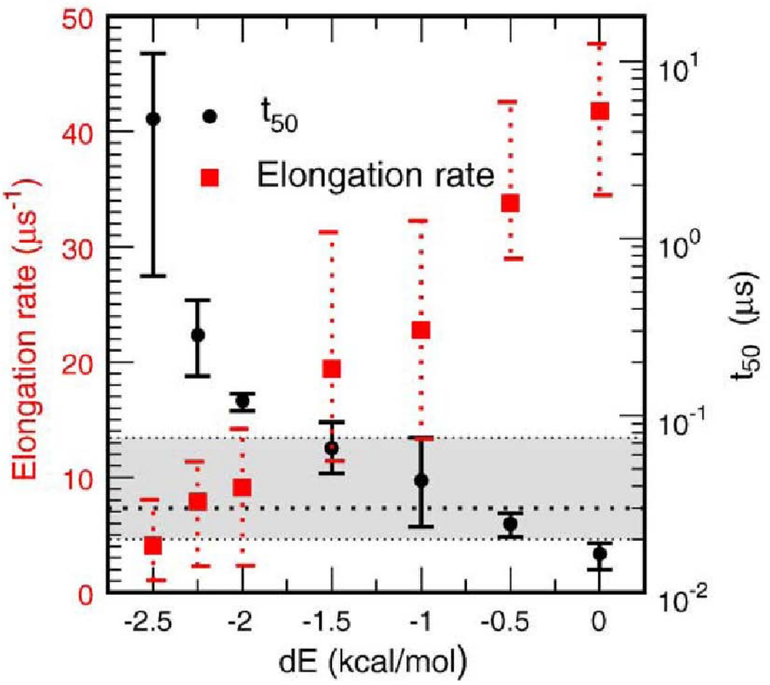


FIG. 6. Influence of the amyloidogenicity parameter dE on the kinetics of the CGF model. The time needed to reach 50% of the maximal amplitude t_{50} (black circles and y-axis legend on the right) and the elongation rate (red squares and y-axis legend on the left) are displayed for seven dE values. Symbols represent the average value of ten independent runs, and the error bars are the maximum and minimum values. The broken line and the gray band indicate the average and the maxmin values for the time of micelle formation, respectively. Reprinted from[37] with permission by Elsevier.

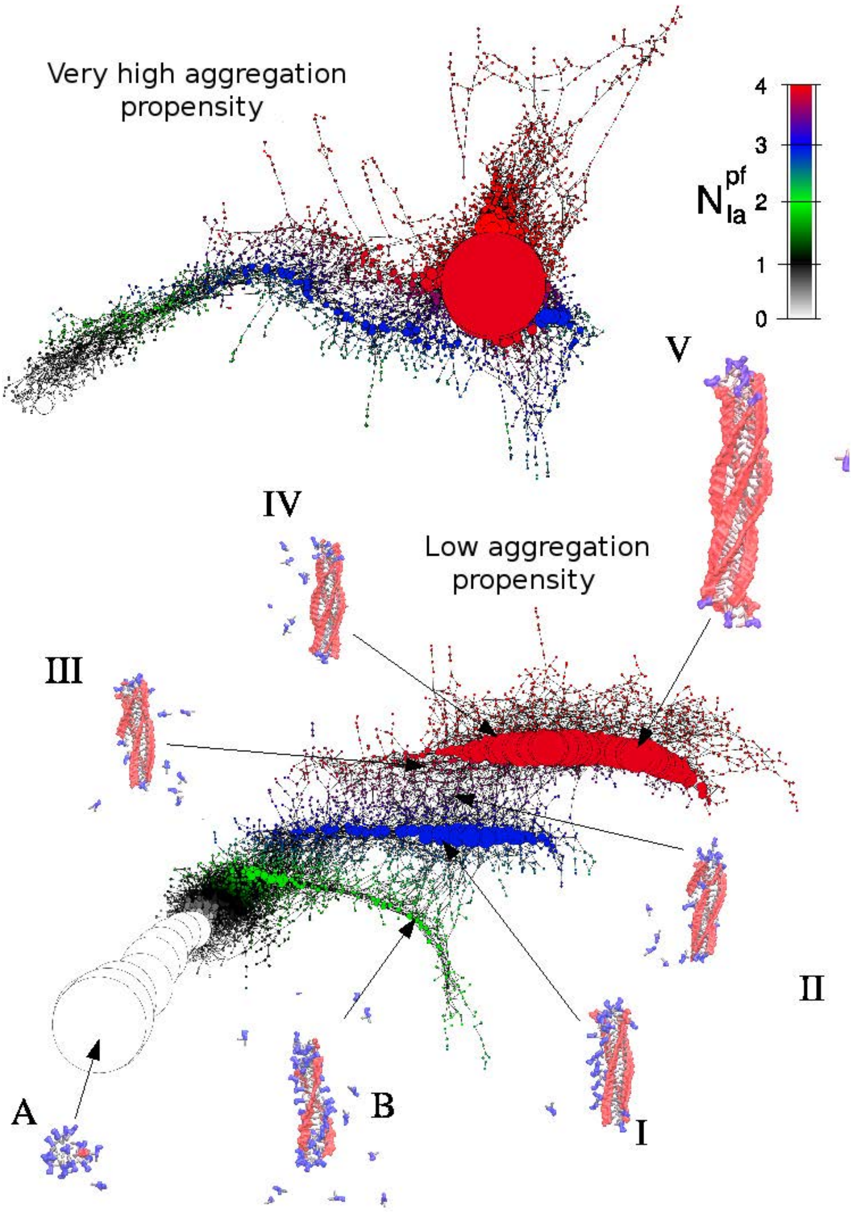


FIG. 7. Aggregation state network. The size of the largest aggregate N_{la} and its number of protofilaments N_{la}^{pf} were used to cluster all simulation snapshots into states (i.e., nodes of the network). The size and the colour of nodes correspond to the statistical weight and the number of protofilaments N_{la}^{pf}, respectively. Links are direct transitions within 0.5 ns of Langevin dynamics. Note the much higher heterogeneity of protofibrillar intermediates for the β-unstable (dE = -2.5 kcal/mol, bottom) as compared to the β-stable (dE = -1.5 kcal/mol, top) model. The insets show (proto)fibrillar structures that are representative of each region of the aggregation state network. In these structures, monomers in the amyloid-competent conformer β and amyloid-protected conformer π are in red and blue, respectively. Furthermore, hydrophobic spheres are gray and hydrophlic spheres are not shown for visual clarity. Reprinted from[50] with permission by Elsevier.

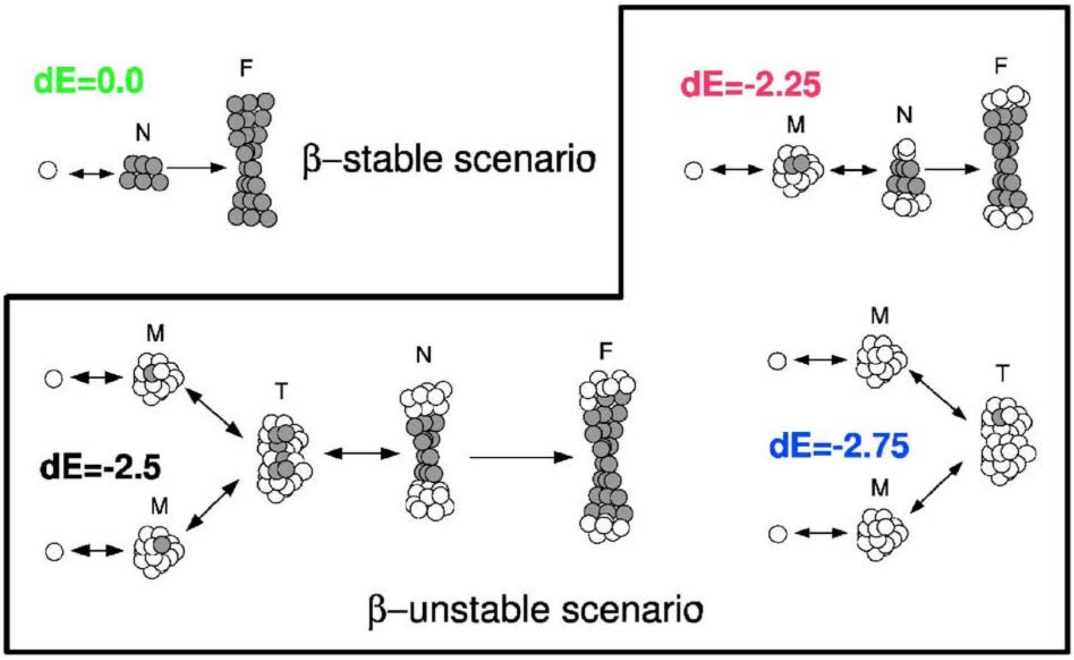


FIG. 8. Observed nucleation scenarios of the CGF peptide model. Black and white circles represent the amyloid-competent conformer β and amyloid-protected conformer π, respectively. CGF peptides with high values of amyloidogenic tendency nucleate without intermediates, while poorly amyloidogenic CGF peptides can nucleate either through micelle-sized oligomers ($dE =$ -2.25 kcal/mol) or transient oligomers larger than a micelle ($dE =$ -2.5 kcal/mol). A further stabilization ($dE =$ -2.75 kcal/mol) of the protected state prevents fibril formation within the simulation time of about 20 μs. M, micelle; N, nucleus; T, transient oligomer; F, fibril. Reprinted from[37] with permission by Elsevier.

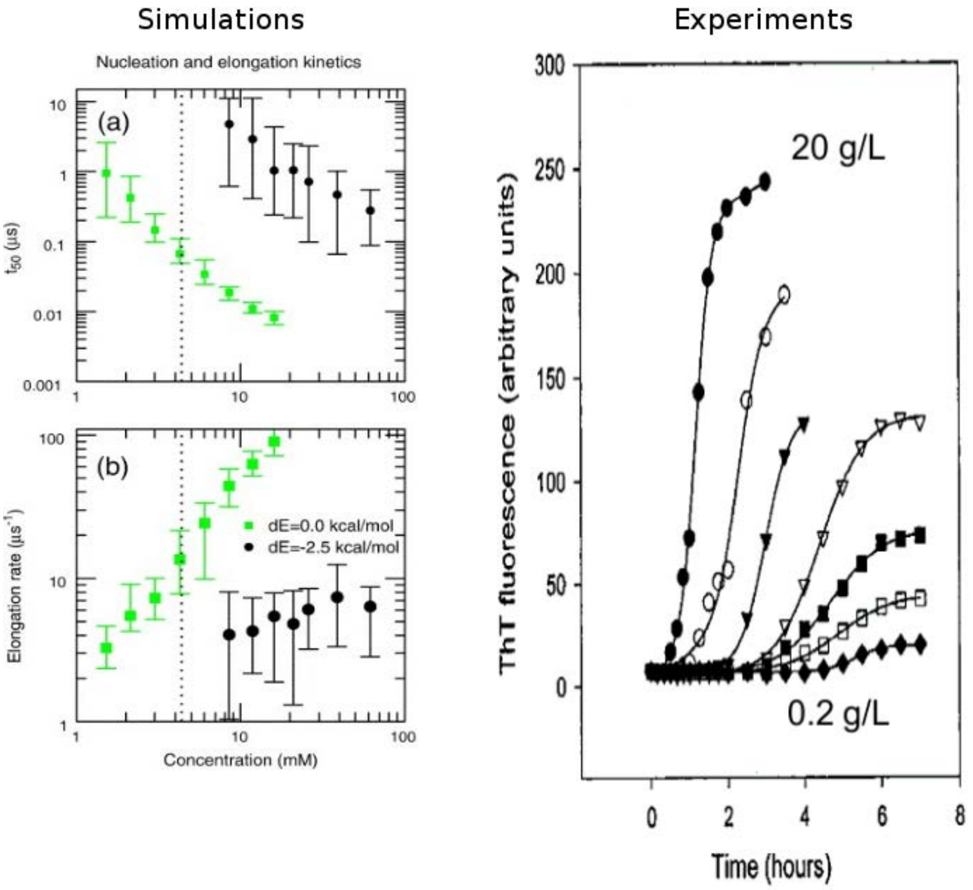


FIG. 9. Influence of peptide concentration on aggregation kinetics. (Left) Effect of concentration on the lag phase time t_{50} (a) and elongation rate (b) for low and high values (dE = -2.5 kcal/mol, black circles; dE = 0.0 kcal/mol, green squares, respectively) of the amyloidogenic tendency. The symbols represent the average value calculated from 15 simulations for dE = -2.5 kcal/mol and ten simulations for dE = 0.0 kcal/mol. The error bars represent the minimum and the maximum values. The vertical dotted line indicates the critical concentration of micelle formation. (Right) Influence of the initial monomeric concentration on the kinetics of insulin fibril formation as measured by Thioflavin T fluorescence[58]. Note that the higher the concentration of monomeric insulin is at the beginning of the experiments, the shorter is the lag phase and the faster is the elongation rate. Reprinted from[37] (left) and[58] (right) with permission by Elsevier (left) and American Chemical Society (right).

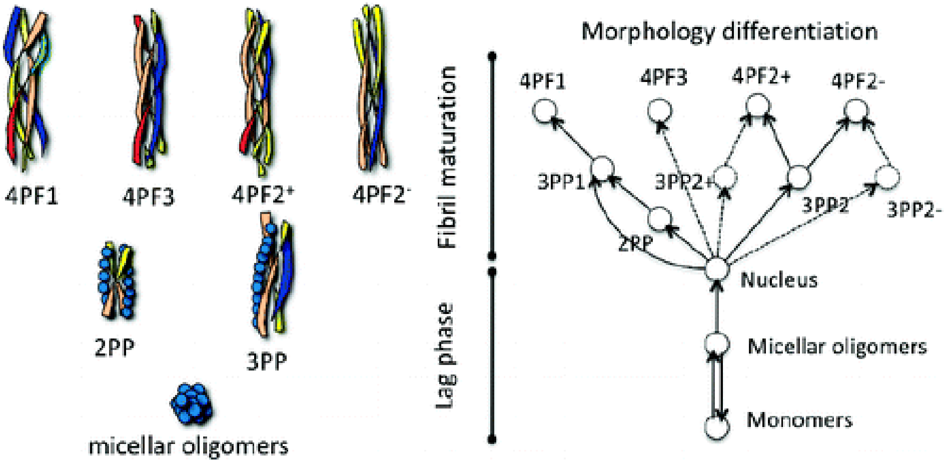


FIG. 10. Morphology differentiation and kinetic control of fibril polymorphism. (Left) Morphologies of mature fibrils and prefibrillar species. (Top) Mature fibrils display a 4-protofilament structure (4PF). The 4PF morphologies have different orientation of the protofilaments, organization of up and down protofilaments, and thickness of the fibril. (Bottom) The prefibrillar species are: the micellar oligomers M, consisting of π-monomers (blue beads) aggregated through hydrophobic forces; the 2-protofilament protofibril (2PP), and the 3-protofilament protofibril (3PP), which are early stages of fibril maturation, where the π-monomers are deposited onto the lateral surface of the fibril, and the β-monomers make up the protofilaments (colored ribbons). (Right) Branched tree illustration of the morphology differentiation process as observed in the simulations. Reprinted from[64] with permission by American Chemical Society.

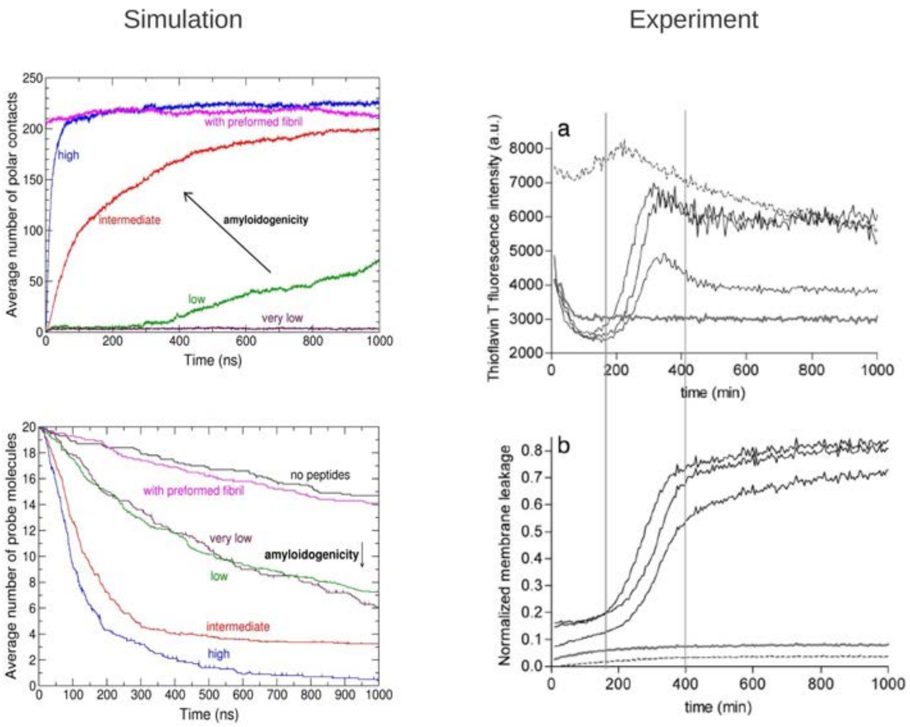


FIG. 11. Comparison of CGF model simulations and experimental data on fibril formation in the presence of lipid vesicles. (Left) Simulation results. (Left, top) Influence of peptide amyloidogenicity on fibril growth kinetics and vesicle leakage. A single parameter, the energy difference between amyloid-competent and amyloid-protected conformations of the peptides, is varied in different simulations to tune amyloidogenicity. Time series of the average number of ordered polar contacts between monomers (corresponding to the degree of fibrillation). (Left, bottom) Average number of probes inside the vesicle, in the absence (black) or presence (colors) of peptides, and for simulations where a preformed fibril was used instead of dispersed peptides. (Right) Experimental data: Effect of human islet amyloid polypeptide (hIAPP) fibril growth on membrane leakage[69]. Thioflavin T fluorescence intensity (Right, top) and induced membrane leakage (Right, bottom) of three hIAPP samples (black curves), together with representative traces for mouse IAPP variant which is known to be non-toxic amyloid polypeptide (gray lines) and preformed hIAPP fibrils (dashed lines) are shown. The two vertical lines are shown to facilitate comparison of the kinetic traces in top and bottom panel. Reprinted from[73] (left) and[69] (right) with permission by Elsevier and by Copyright 2008 National Academy of Sciences, U.S.A., respectively.

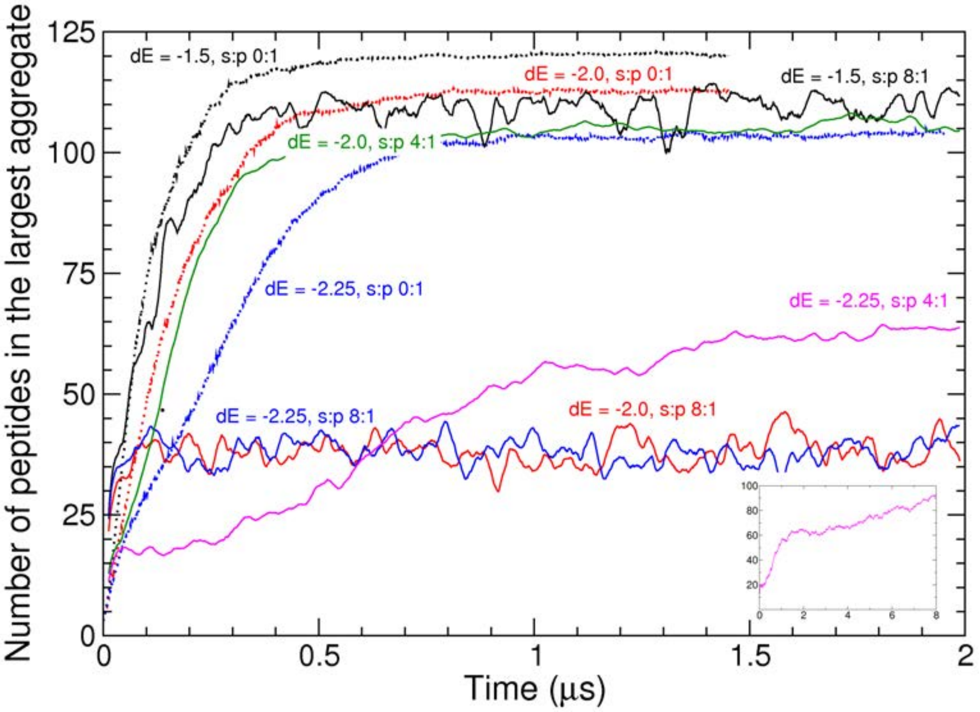


FIG. 12. Amyloid aggregation in the presence and absence of surfactants. The number of peptides in the largest aggregate is averaged over 20 runs at each simulation condition, i.e., for each value of aggregation propensity (dE) and each surfactant/peptide concentration ratio (s:p). Free peptides display fast aggregation without any noticeable lag phase. At surfactant:peptide ratio of 4:1, peptides aggregate into fibrils, but aggregation is much slower for peptides with low amyloidogenicity (see the inset for $dE =$ -2.25 and surfactant:peptide ratio of 4:1, extended to 8 μs). At a surfactant:peptide ratio of 8:1, aggregation is completely inhibited for $dE =$ -2.0 and -2.25 kcal/mol, while highly amyloidogenic peptides ($dE =$ -1.5 kcal/mol) are barely affected.

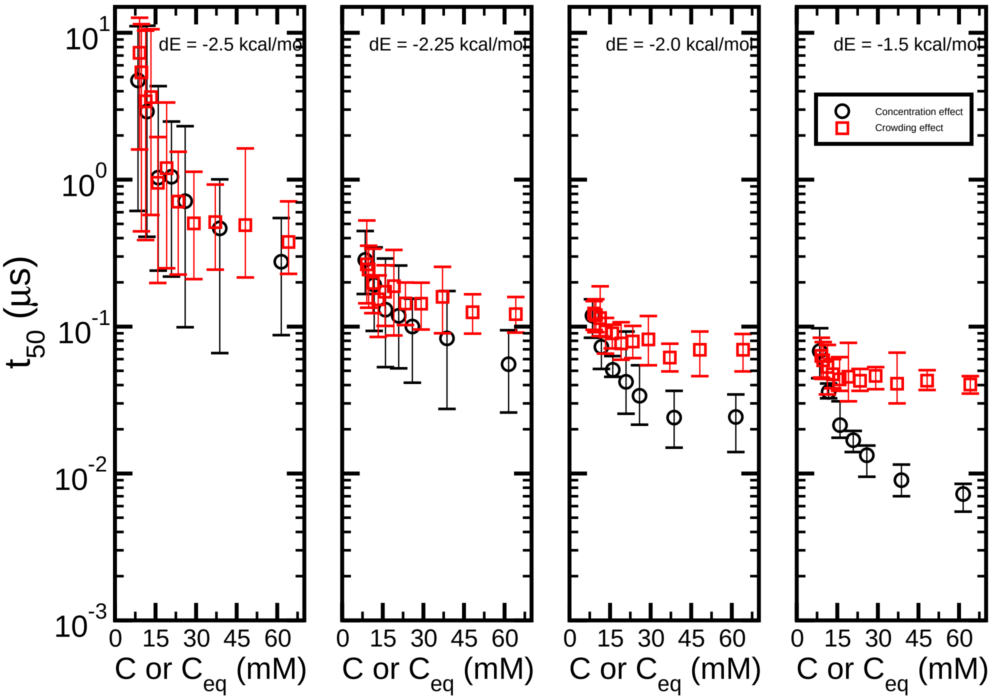


FIG. 13. The deviation of aggregation kinetics between crowded and bulk conditions depends on the amyloidogenicity. The time t_{50}, at which the growing fibril has reached 50% of the polar contacts of the mature fibril, is shown as a function of concentration. Black circles are t_{50} values calculated at different peptide concentrations in the absence of crowders, while red squares are t_{50} values at different equivalent concentrations C_{eq} obtained by varying the number of crowders. Symbols represent the average value of 10 independent runs and the error bars are the minimum and maximum values. Reprinted from[81] with permission by American Chemical Society.

Chapter 3

Crowding effects on amyloid aggregation kinetics

Magno, A., Caflisch, A. and Pellarin, R. *Jour. Phys. Chem. Lett.*, **2010**, *1, 3027-3032*

The Journal of
Physical Chemistry
Letters

pubs.acs.org/JPCL

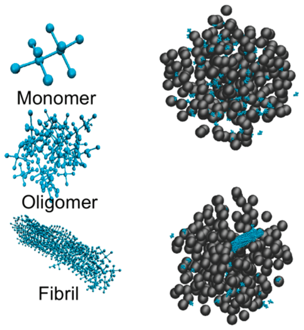
Crowding Effects on Amyloid Aggregation Kinetics

Andrea Magno, Amedeo Caflisch,* and Riccardo Pellarin*

Department of Biochemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

ABSTRACT Biological protein self-assembly occurs in the cellular milieu, densely occupied by other macromolecules which do not participate directly in the aggregation process. Excluded volume effects arising in such a crowded environment deeply affect the thermodynamics and kinetics of biological processes, like protein folding, ligand binding, and protein aggregation. Here, Langevin dynamics simulations of a simplified model of an amphipathic polypeptide are used to investigate how macromolecular crowding influences the amyloid aggregation kinetics. The simulations show that the net influence of macromolecular crowding on the self-assembly process is the result of two competing effects: oligomer stabilization and solution viscosity increase. Notably, the net effect crucially depends on the aggregation propensity and pathways. Therefore, comparative studies of concentration and crowding effects on the kinetics of amyloid aggregation could shed light on the underlying self-assembly mechanism.

SECTION Biophysical Chemistry



Amyloid fibrils are ordered polypeptide aggregates that have been related to several neurodegenerative pathologies, such as Alzheimer's, Parkinson's, Huntington's and prion diseases,[1,2] and, more recently, also to biological functionalities.[3,4] These findings have paved the way to a wide range of experimental and computational studies aimed at understanding the details of the fibril formation mechanism.

Most of these investigations were usually performed in ideal homogeneous conditions, though the actual cellular milieu is a much more complex environment. Several studies have pointed out the effects of geometric confinement on fibril formation and protein folding.[5,6] A universal property of the cells is that they are crowded.[7,8] Indeed, it has been estimated that 20−30 % of the cell cytoplasm is occupied by proteins, RNA, membranes, polysaccharides, and several organelles.[9] Although the concentration of every species is low, these macromolecules exclude a significant fraction of the total available volume.[10] The nonspecific excluded volume effect is expected to sensibly affect all biological reactions in which proteins are involved.[11] Using scaled particle theory,[12] Minton and Ellis have predicted that macromolecular crowding dramatically increases the association rate of proteins and the relative stability of the unfolded and native state.[13,14] Computational studies have shown that the presence of inert, repulsive cosolutes stabilizes the native, compact state of proteins[15,16] and that the presence of nanoparticles is able to catalyze amyloid aggregation.[17] Experiments[18−21] conducted with large and weakly interacting macromolecules such as polyethyleneglycol (PEG) and Fycoll have confirmed these theoretical predictions, showing that the excluded volume effect is able to modify the subtle equilibrium between the folded, functional state and aberrant structures prone to aggregation.

Earlier, a very simple coarse-grained model of an aggregation-prone, amphipathic peptide was developed to investigate the kinetics of ordered aggregation.[22] The peptide monomer has a single degree of freedom, and the relative free-energy profile has only two minima, corresponding to the aggregation-prone and aggregation-protected states. By varying a single parameter of the model, that is, by reducing the β-aggregation propensity, the roughness of the free-energy landscape and the heterogeneity of the fibril elongation pathway increase. In previous simulation studies, heterogeneous kinetics of aggregation and multiple pathways were observed in bulk solution and in the presence of lipid bilayers. At high amyloidogenic conditions, the process of fibril formation is downhill and fast, whereas at low amyloidogenic conditions, several intermediates are detected, and the nucleation occurs through a micellar oligomer.[23] It has also been pointed out that amyloidogenicity determines the effect on peptide aggregation of the presence of lipid bilayer; while aggregation-prone peptides fibrillate faster by adsorbing on the bilayer surface, the ordered self-assembly of poorly amyloidogenic peptides is hindered by the vesicles.[24,25] In this work, a spherical model of softly repulsive crowders is used, together with the simple model of the amphipathic peptide, to investigate amyloid aggregation kinetics at different concentrations of crowders. It is found that the influence of the crowders has a pronounced dependence on the amyloidogenic tendency of the peptide.

The simulations were performed with 125 peptides and a number of softly repulsive crowders ranging from 250 to 5000 (see Methods and Table 1). Snapshots of the simulation

Received Date: July 16, 2010
Accepted Date: September 12, 2010

Table 1. Excluded Volume Fraction (ϕ), Equivalent Concentration (C_{eq} = 8.5 mM /(1 − ϕ)), and Peptide Self-Diffusion Coefficient ($\mathcal{D}$) at Different Crowder Contents n_C[a]

n_C	ϕ	C_{eq} (mM)	$\mathcal{D}$ (Å^2/ps)
250	0.066	9.10	1835
500	0.13	9.77	1765
1000	0.25	11.3	1435
1500	0.36	13.3	1234
2000	0.46	15.8	1075
2500	0.56	19.1	910
3000	0.64	23.4	779
3500	0.71	29.2	686
4000	0.77	37.1	617
4500	0.82	48.2	547
5000	0.87	64.2	498

[a] See the Methods section for details.

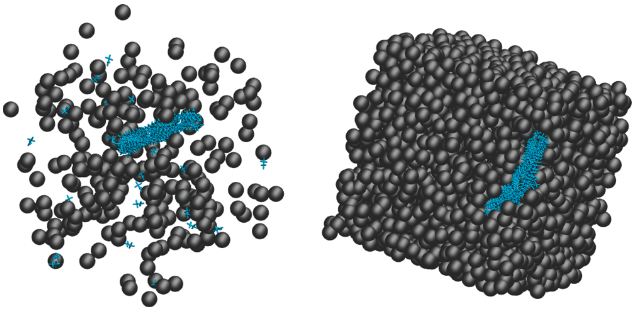


Figure 1. Snapshots of simulation systems at low (n_C = 250, left) and high crowder concentrations (n_C = 3000, right). Peptides are shown in a stick model in cyan, while crowders are black spheres.

system with a low (n_C = 250) and high (n_C = 3000) number of crowders are shown in Figure 1.

To compare crowding effects with an increase in peptides concentration, simulations in the absence of crowders with 125 peptides and a progressively decreasing volume of the simulation box (yielding peptide concentrations ranging from 8.5 to 61.5 mM) were also carried out. For each of these conditions, multiple runs were started (with different initial velocities) for each of four values of the amyloidogenicity (i.e., dE = −2.5, − 2.25, − 2.0, and −1.5 kcal/mol; see Methods for details) to investigate the dependence on the aggregation propensity. The temperature in all simulations was 310 K.

Crowders Effect on Lag Phase. Supercritical oligomers are assemblies whose size is larger than the nucleus, where the nucleus is defined as the oligomer that has a 50% probability to form a fibril[22] (see Methods and Table S1 (Supporting Information) for details). To specifically assess how macromolecular crowding influences the nucleation kinetics, the number of supercritical oligomers has been calculated as a function of time along each simulation (Figure 2). During the lag phase, the number of oligomers of supercritical size is roughly zero, that is, only small oligomers or monomers are present in the system, while for long time, only one aggregate, which corresponds to the mature fibril, is present. At intermediate times, the number of oligomers of supercritical size can be higher than one when two or more elongating fibrils are present simultaneously in the simulation box. To directly compare the effect of peptide concentration and macromolecular crowding, the equivalent concentration C_{eq}, defined as the ratio between the number of peptides and the volume unoccupied by crowders (see Methods), is introduced. Concentration increment and macromolecular crowding have sensibly different effects. They both increase the average number of supercritical oligomers at intermediate times, but the concentration also accelerates the nucleation step (corresponding to the shift of the peak toward shorter times in the time series in Figure 2, left), while the crowding has a much less pronounced influence on the nucleation step because the decrease in peptide diffusivity limits monomer encounters.

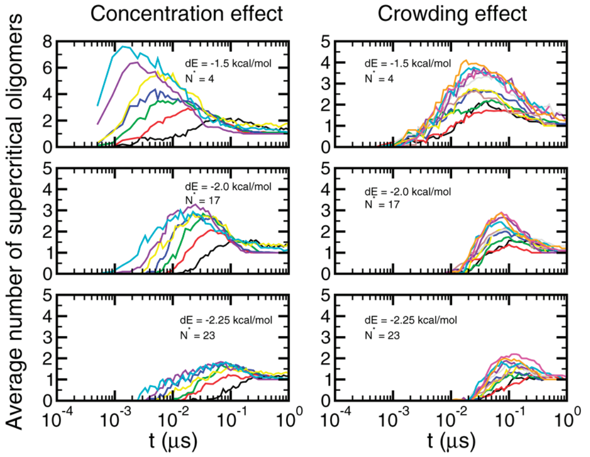


Figure 2. Peptide concentration but not crowder content accelerates the nucleation step. Average number of supercritical oligomers (i.e., with size larger than the nucleus aggregation number N^*) as a function of simulation time. Each curve is an average over 10 runs. The size N^* of the nucleus increases upon destabilization of the amyloid prone state (see Table S1, Supporting Information). (Left panels) Different colors represent different peptide concentrations, ranging from (black) C = 8.5 to (cyan) 61.5 mM. Note that the y-axis range is different for the top plot. (Right panels) Different colors represent different crowder contents, ranging from (black) n_C = 250, C_{eq} = 9.1 to (orange) n_C = 5000, C_{eq} = 64.2.

The mean first appearance of a supercritical oligomer t_{N^*} reports on the length of the lag phase. Crowding and concentration effects yield similar lag phases only for the peptide model with the lowest amyloidogenicity (Figure 3, left). On the contrary, for the more amyloidogenic peptide models, the lag phase becomes shorter by raising the concentration while it is only marginally influenced by the crowder content (Figure 3, right).

Crowders Effect on Oligomers Stability. To interpret the kinetics at low dE, it is useful to investigate the stability of oligomers that are competent to fibril nucleation and are explored during the lag phase. The presence of a large amount of inert macromolecules that exclude volume to peptides favors association.[14] Since oligomers are more compact than an equivalent amount of dispersed monomers, they are thermodynamically stabilized by an increase of excluded volume. Therefore, to study the thermodynamic properties of the metastable oligomers that appear during the lag phase, multiple simulations were carried out at different crowder

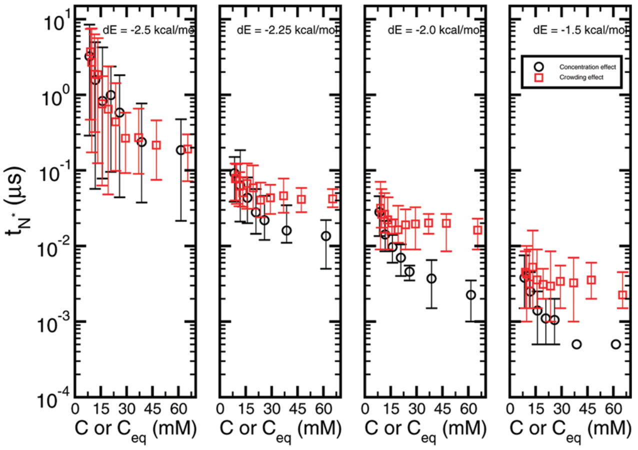


Figure 3. Time of appearance of the first supercritical oligomer t_{N^*}. Black circles are t_{N^*} values calculated at different peptide concentrations without crowders, whereas red squares are t_{N^*} values at different equivalent concentrations C_{eq} obtained by varying the number of crowders (see Table 1). Symbols represent the average value of 10 independent runs, and the error bars are the minimum and maximum values. Note that the two data points at highest peptide concentration (black circles) in the right panel do not have an error bar as the first supercritical oligomer appeared within the first coordinate saving interval of 0.5 ns.

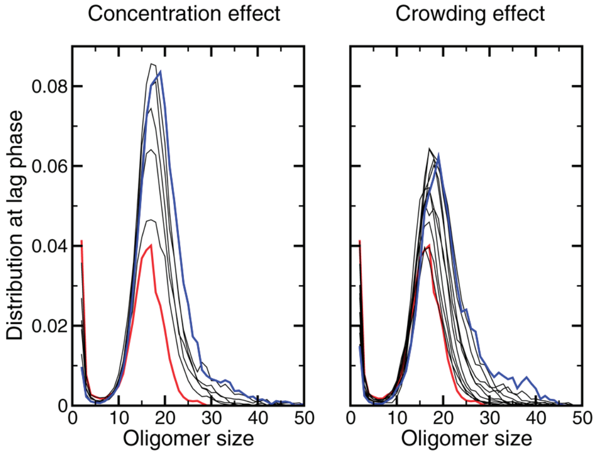


Figure 4. The increase in oligomer size is more pronounced for higher peptide concentration than that for higher crowding content. Oligomer size distribution, evaluated for dE = −3.5 kcal/mol, at different values of the crowder content and peptide concentration. The red curve in each plot corresponds to C = 8.5 mM (n_C = 0), while the blue curves correspond to the highest values of concentration (C = 61.5 mM and C_{eq} = 64.2 mM in the left and right panel, respectively). The frequency of the isolated monomers ranges from 0.64 to 0.12 and 0.64 to 0.31 on the left and right plots, respectively, and is not shown to avoid compression of the oligomeric peak.

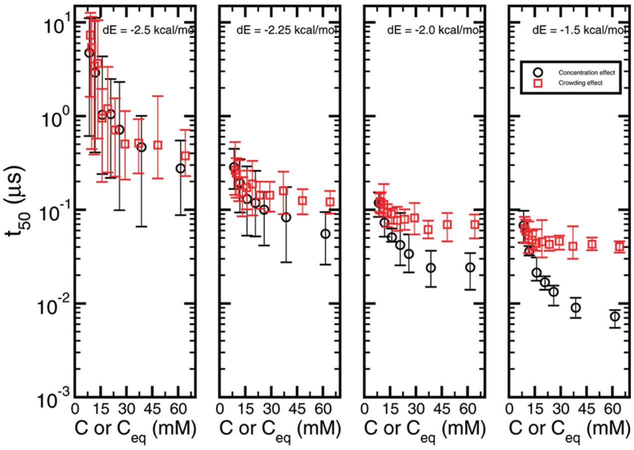


Figure 5. Differences in aggregation kinetics upon raising the peptide concentration or crowder content depend on amyloidogenicity. The time t_{50} at which the growing fibril has reached 50% of the polar contacts of the mature fibril is shown as a function of concentration. Black circles are t_{50} values calculated at different peptide concentrations in the absence of crowders, while red squares are t_{50} values at different equivalent concentrations C_{eq} obtained by varying the number of crowders (see Table 1). Symbols represent the average value of 10 independent runs, and the error bars are the minimum and maximum values.

concentrations with 125 peptides with a stable aggregation-protected state (dE = −3.5 kcal/mol). This value has been chosen to inhibit the fibril nucleation but not the oligomer formation. The distribution of oligomer sizes at different crowder concentrations shows that larger oligomers are stabilized by the excluded volume effect, which is proportional to the amount of crowders (see Figure 4). Therefore, macromolecular crowding stabilizes oligomers, but it does so less than an equivalent increase of peptide concentration. It has been shown experimentally that oligomeric intermediates can be stabilized significantly by macromolecular crowding.[19] In this context, it is likely that if off-pathway oligomers are present in the aggregation mechanism, the aggregation kinetics will be slowed down by crowding. Such an effect is not observed because off-pathway oligomers cannot be reproduced by the simple model used in this simulation study.

Crowders Effect on Fibril Formation Kinetics. Macromolecular crowding is more effective for the less aggregating peptides (dE = −2.5 kcal/mol), for which the time needed to establish half of the number of polar contacts present in the mature fibril (t_{50}) diminishes by an order of magnitude at high excluded volume fractions (Figure 5). This trend is analogous to that observed experimentally by Munishkina et al., who have studied the effect of increasing the PEG concentration on the α-synuclein aggregation process.[18] Note that the time t_{N^*} (shown in Figure 3) reports on the lag phase, while t_{50} is a mixed measure of both the lag phase and the inverse rate of elongation. Specifically, the t_{50} is dominated by the lag phase for low amyloidogenic peptide models because the nucleus size contains almost as many monomers as half of the final fibril. In contrast, it mainly reflects the inverse rate of elongation for high amyloidogenicity potentials whose nucleus size is very small.

The effect of macromolecular crowding is less pronounced for the more aggregating peptides, for which the t_{50} values reach a plateau already at low crowder content. Interestingly, for the peptide models with low aggregation propensity (dE = −2.5, − 2.25 kcal/mol), the effect of macromolecular crowding and peptide concentration is similar (Figure 5, left). In contrast, for peptides more prone to aggregation (dE = −2.0, − 1.5 kcal/mol), the acceleration of self-assembly promoted by the crowders is much smaller than that caused by an equivalent increase in peptide concentration (Figure 5, right). Moreover, this discrepancy grows by increasing the aggregation propensity (i.e., at higher dE) and is similar to the one previously observed in the trend of the t_{N^*} for different values of dE (see Figure 3). The different effects of crowders and peptide concentration have not been observed previously. It would be interesting to validate them by kinetic measurements on peptides that aggregate very fast like diphenylalanine.[26]

Two Main Scenarios of Aggregation. The aggregation kinetics of a simple model of an amphipathic peptide have been investigated at different concentrations of inert crowders. The excluded volume effect stabilizes larger oligomers (Figures 2 and 4). However, at high crowder content, the solution becomes more viscous (Figure S1, Supporting Information), and the peptide mobility decreases because each monomer is locally confined by the crowders (see Table 1). The confinement decreases the rate of monomer encounters and, therefore, association (Figures 3 and 5). Two main scenarios emerge from the present simulation study.

For peptides with low aggregation propensity, the self-association process is transition-state-limited, where the kinetic bottleneck is the formation of the fibril nucleus. In this case, since the oligomers, including the nucleus, are thermodynamically favored (with respect to the isolated monomers) by the excluded volume effect, macromolecular crowding accelerates peptide assembly and has an effect analogous to that of an increase in peptide concentration (Figure 5, left).

On the other hand, when the aggregation mechanism is fast and proceeds directly from monomers to the fibril, the process is diffusion-limited, and the thermodynamic stabilization of oligomers is less important than the reduction in peptide mobility. In this case, the bottleneck is not the formation of the nucleus; the rate-limiting step for peptides that show a direct aggregation mechanism is the elongation of the fibril. Therefore, in this case, macromolecular crowding is much less efficient in accelerating peptide self-association than an equivalent increase of the peptide concentration since the peptides motion is hindered by the crowders (Figure 5, right). The very different relative influence of crowders and concentration observed in the simulations is a new finding which suggests that kinetic experiments at different crowder contents might discriminate between different mechanisms of aggregation, that is, downhill from barrier-limited.

METHODS

Peptide Model. The amphipathic peptide model adopted here has been described in detail elsewhere.[22] Briefly, each peptide monomer consists of 10 beads. Four of the beads carry partial charges of $\pm 0.4e$, thereby generating two dipoles. Four of the uncharged spheres are hydrophilic, and the other two are hydrophobic. Interactions between monomers depend on van der Waals and electrostatic forces. The former approximate both steric and hydrophobic effects, while the latter are specific dipole−dipole interactions responsible for the ordered stacking of monomers incorporated in the fibril. A parallel polar contact is formed whenever two charged beads of different monomers are closer than 5 Å. Different aggregation pathways are obtained by changing the relative stability of the amyloid-competent (β) and amyloid-protected (π) states, which are ruled by the dihedral energy difference of the single rotatable bond of the monomer. The energy difference of these two states, d$E = E_\pi - E_\beta$, can be therefore interpreted as the β-aggregation propensity of the polypeptide. Since the peptide has only one degree of freedom, dE is close to the free-energy difference between the two aforementioned states. For instance, when dE = 0 kcal/mol, the π and β states are equally populated, whereas for dE = −1.5 and −2.5 kcal/mol, the π state is about 15 and 100 times more populated than the β state, respectively.

Crowder Model. The crowders are neutral spherical particles that interact with the peptides and with each other by means of the Lennard-Jones potential, given by the following equation

$$V_{ij} = \epsilon_{ij}^{\min}\left[\left(\frac{R_{ij}^{\min}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{ij}^{\min}}{r_{ij}}\right)^{6}\right] \qquad (1)$$

where $R_{ij}^{\min}$ is the distance at the minimum of the potential, or the van der Waals radius of the molecule, $\epsilon_{ij}^{\min}$ is the depth of the potential well, and r_{ij} is the interparticle distance. The value of $R_{ij}^{\min}$ is set to 7.5 Å, and ϵ = 0.1 kcal/mol for the crowders, which therefore behave as impenetrable soft spheres. The mass of the crowders has been set to 6.5 kDa. The interactions between the crowder and peptide are also modeled with a Lennard-Jones potential. In this case, the $\epsilon_{ij}^{\min}$ parameter is set to 0.01 kcal/mol to make the crowders softly repulsive particles; the optimal distance is obtained by using the arithmetic mean.

Simulation Protocol with Crowders. Each simulated system is made up of 125 peptides in a cubic box of size equal to 290 Å, at a peptide concentration of 8.5 mM and a variable number of crowders (see Table 1). Periodic boundary conditions are imposed. All of the simulations were carried out at 310 K by means of Langevin dynamics using CHARMM.[27,28] Since aggregation is a stochastic process, 10 simulations with different starting velocities were run for every concentration of crowders and for each value of the β-aggregation propensity of the peptides. The time interval for saving the coordinates of the system was 0.5 ns. It is important to note that the relative populations of β and π states of the isolated monomer are influenced solely by the value of dE and not by the crowder content because the interactions between crowders and peptides are not sensitive to the two different states of the monomers.

Simulation Protocol without Crowders. These simulations were carried out as described above (i.e., 125 peptides and 310 K) in the absence of crowders and using box sizes ranging from 150 to 290 Å, corresponding to a peptide concentration ranging from 61.5 to 8.5 mM, respectively.

Excluded Volume. A rigorous protocol to calculate the volume fraction that the crowders exclude to peptides at different concentrations has been established. According to the linear approximation, the excluded volume fraction is given by the following equation

$$\phi^{l}(n_{\mathrm{C}}) = \frac{V_{\mathrm{excl}}}{V_{\mathrm{tot}}} = \frac{1}{V_{\mathrm{tot}}}\frac{4}{3}\pi R_{\mathrm{eff}}^{3} n_{\mathrm{C}}$$

where $R_{\mathrm{eff}} = R_{\mathrm{C}} + R_{\mathrm{P}}$ is the sum of the radius of the crowder R_{C} and the radius of the peptide R_{P} and n_{C} is the number of crowders in the total volume. This approximation holds only if the crowders are significantly larger than the peptide, that is, when $R_{\mathrm{eff}} \approx R_{\mathrm{C}}$. If the radius of the peptides is non-negligible, the excluded volume must include two corrections arising from the presence of crowders whose mutual distance is lower than $2R_{\mathrm{eff}}$ (see Figure S2, Supporting Information).

Since the peptide is not fully spherical, a heuristic evaluation of R_{eff} is needed. A simulation of an "ideal" solution mixture of 125 peptides and 500 crowders was run, where the polar interactions of the peptides were switched off, as well as the hydrophobic ones, to prevent them from aggregating. From a 500 ns run, the radial distribution functions $g(r)$ for each pair of species in the solution were calculated. The effective radius R_{eff} has been chosen as the radius where the radial distribution function $g_{\mathrm{CP}}(r)$ between the crowders and the peptides has the value of 0.5 (cfr. Figure S3, Supporitng Information), obtaining $R_{\mathrm{eff}} = 11.6$ Å. The effective excluded volume fraction ϕ is calculated by dividing the whole simulation box in "infinitesimal" cubic elements and counting the number of cubes whose distance from any crowder particles center is less than R_{eff}, which is the minimal approachable distance between a crowder and a peptide.

Simulations of 500 ns with different numbers of crowders were used to produce snapshots of the system with different arrangements of the crowders. The excluded volume fraction ϕ at different crowder content n_{C}, reported in Table 1, is calculated as an average over all snapshots of the simulations. The equivalent concentration $C_{\mathrm{eq}} = 8.5\ \mathrm{mM}/(1-\phi)$ is the concentration of peptides in the accessible volume.

Dimension of the Fibril Nucleus. The nucleus is defined as the oligomer that has 50% probability to form a fibril.[22] The crucial feature of such an oligomer is that it is not only formed by a certain number of monomers but that it also owns the sufficient number of monomers in the β state to permit the elongation of the fibril (Table S1, Supporting Information).

Self-Diffusion Coefficient. The self-diffusion coefficient $\mathcal{D}$, which measures the mobility of the isolated peptides, can be calculated through the Einstein relation

$$\lim_{t \to \infty} \langle r^{2}(t) \rangle = 6\mathcal{D}t$$

where $\langle r^{2}(t) \rangle$ is the mean-square displacement. Simulations of 100 ns with different concentrations of crowders and 125 peptides were performed to calculate the mean-square displacement functions. The polar and the hydrophobic interactions of the peptides were switched off to prevent aggregation events and calculate the diffusivity of the single monomers. The self-diffusion coefficients of peptides at different crowder contents $\mathcal{D}(n_{\mathrm{C}})$ (Table 1) were then derived by a linear fit of the long-time behavior of the mean-square displacement (see Figure S4, Supporting Information).

SUPPORTING INFORMATION AVAILABLE Supplementary Figures 1–4 and Table 1. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author:

*To whom correspondence should be addressed. E-mail: caflisch@bioc.uzh.ch (A.C.); pellarin@bioc.uzh.ch (R.P.).

ACKNOWLEDGMENT This work was supported by a grant of the Swiss National Science Foundation to A.C. The simulations were carried out on the Schrödinger compute cluster of the University of Zürich.

REFERENCES

(1) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003,** *426*, 884–890.

(2) Lansbury, P. T.; Lashuel, H. A. A Century-Old Debate on Protein Aggregation and Neurodegeneration Enters the Clinic. *Nature* **2006,** *443*, 774–779.

(3) Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W. Functional Amyloid — From Bacteria to Humans. *Trends Biochem. Sci.* **2007,** *32*, 217–224.

(4) Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D.; et al. Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* **2009,** *325*, 328–332.

(5) Kraineva, J.; Smirnovas, V.; Winter, R. Effects of Lipid Confinement on Insulin Stability and Amyloid Formation. *Langmuir* **2007,** *23*, 7118–7126.

(6) Mittal, J.; Best, R. B. Thermodynamics and Kinetics of Protein Folding under Confinement. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105*, 20233–20238.

(7) Fulton, A. B. How Crowded Is the Cytoplasm. *Cell* **1982,** *30*, 345–347.

(8) Zimmerman, S. B.; Trach, S. O. Estimation of Macromolecule Concentrations and Excluded Volume Effects for the Cytoplasm of Escherichia-Coli. *J. Mol. Biol.* **1991,** *222*, 599–620.

(9) Ellis, J. R. Macromolecular Crowding: Obvious but Underappreciated. *Trends Biochem. Sci.* **2001,** *26*, 597–604.

(10) Minton, A. P. The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media. *J. Biol. Chem.* **2001,** *276*, 10577–10580.

(11) Zhou, H. -X.; Rivas, G.; Minton, A. P. Macromolecular Crowding and Confinement: Biochemical, Biophysical, And Potential Physiological Consequences. *Ann. Rev. Biophys.* **2008,** *37*, 375–397.

(12) Lebowitz, J. L.; Rowlinson, J. S. Thermodynamic Properties of Mixtures of Hard Spheres. *J. Chem. Phys.* **1964,** *41*, 133–138.

(13) Minton, A. P. Models for Excluded Volume Interaction between an Unfolded Protein and Rigid Macromolecular Cosolutes: Macromolecular Crowding and Protein Stability Revisited. *Biophys. J.* **2005,** *88*, 971–985.

(14) Ellis, R. J.; Minton, A. P. Protein Aggregation in Crowded Environments. *Biol. Chem.* **2006,** *387*, 485–497.

(15) Cheung, M. S.; Klimov, D.; Thirumalai, D. Molecular Crowding Enhances Native State Stability and Refolding Rates of Globular Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102*, 4753–4758.

(16) Mittal, J.; Best, R. B. Dependence of Protein Folding Stability and Dynamics on the Density and Composition of Macromolecular Crowders. *Biophys. J.* **2010,** *98*, 315–320.
(17) Auer, S.; Trovato, A.; Vendruscolo, M. A Condensation-Ordering Mechanism in Nanoparticle-Catalyzed Peptide Aggregation. *PLoS Comput. Biol.* **2009,** *5*, 1–7.
(18) Munishkina, L. A.; Cooper, E. M.; Uversky, V. N.; Fink, A. L. The Effect of Macromolecular Crowding on Protein Aggregation and Amyloid Fibril Formation. *J. Mol. Recognit.* **2004,** *17*, 456–464.
(19) Munishkina, L. A.; Ahmad, A.; Fink, A. L.; Uversky, V. N. Guiding Protein Aggregation with Macromolecular Crowding. *Biochemistry* **2008,** *47*, 8993–9006.
(20) Zhou, B. -R.; Zhou, Z.; Hu, Q. -L.; Chen, J.; Liang, Y. Mixed Macromolecular Crowding Inhibits Amyloid Formation of Hen Egg White Lysozyme. *Biochim. Biophys. Acta* **2008,** *1784*, 472–480.
(21) Zhou, Z.; Fan, J. -B.; Zhu, H. -L.; Shewmaker, F.; Yan, X.; Chen, X.; Chen, J.; Xiao, G. -F.; Guo, L.; Liang, Y. Crowded Cell-Like Environment Accelerates the Nucleation Step of Amyloidogenic Protein Misfolding. *J. Biol. Chem.* **2009,** *284*, 30148–30158.
(22) Pellarin, R.; Caflisch, A. Interpreting the Aggregation Kinetics of Amyloid Peptides. *J. Mol. Biol.* **2006,** *360*, 882–892.
(23) Pellarin, R.; Guarnera, E.; Caflisch, A. Pathways and Intermediates of Amyloid Fibril Formation. *J. Mol. Biol.* **2007,** *374*, 917–924.
(24) Friedman, R.; Pellarin, R.; Caflisch, A. Amyloid Aggregation on Lipid Bilayers and Its Impact on Membrane Permeability. *J. Mol. Biol.* **2009,** *387*, 407–415.
(25) Friedman, R.; Pellarin, R.; Caflisch, A. Soluble Protofibrils As Metastable Intermediates in Simulations of Amyloid Fibril Degradation Induced by Lipid Vesicles. *J. Phys. Chem. Lett.* **2010,** *1*, 471–474.
(26) Reches, M.; Gazit, E. Designed Aromatic Homo-Dipeptides: Formation of Ordered Nanostructures and Potential Nanotechnological Applications. *Phys. Biol.* **2006,** *3*, S10–S19.
(27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM — A Program for Macromolecular Energy, Minimization, And Dynamics Calculations. *J. Comput. Chem.* **1983,** *4*, 187–217.
(28) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009,** *30*, 1545–1614.

Crowding Effects on Amyloid Aggregation Kinetics

Supplementary Materials

Andrea Magno[1], Amedeo Caflisch[1*], and Riccardo Pellarin[1*]

[1]*Department of Biochemistry, University of Zürich,*
Winterthurerstrasse 190, CH-8057, Zürich, Switzerland

[*] Authors to whom correspondence should be addressed: caflisch@bioc.uzh.ch, pellarin@bioc.uzh.ch

dE (kcal/mol)	Nucleus size	β-Subdomain size
-2.5	35	18
-2.25	23	10
-2.0	17	17
-1.5	4	4

TABLE 1: The second column lists the number of peptide monomers in the nucleus for different values of the amyloidogenicity dE. The β-subdomain size is the number of monomers in β conformation within the nucleus. The nucleus is defined as the aggregate with 50% probability to further evolve into a fibril and 50% probability to dissolve into isolated monomers.

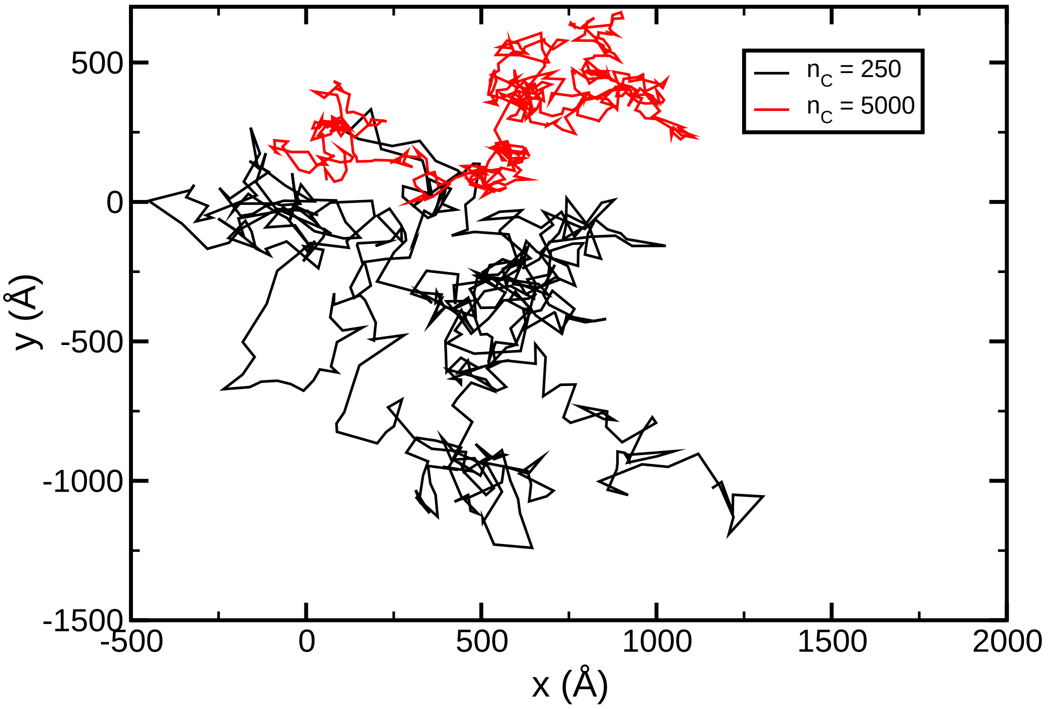


FIG. 1: Projection on the $x-y$ plane of a 20 ns-long trajectory of a peptide at different crowders concentration. The projection shows that at high concentration of crowders the motion of the peptide is hindered.

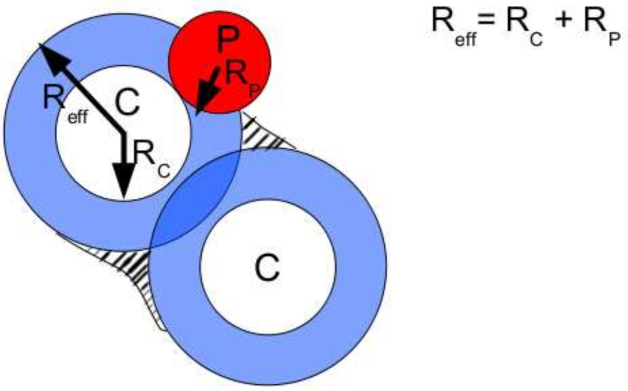


FIG. 2: Effective volume excluded to a peptide (P) by two crowders (C). Since the radius of the peptide is non-negligible with respect to the one of the crowder, the excluded volume evaluation requires two corrections. The first correction accounts for the overlying volume between crowders (darker blue area), and the second derives from the interstitial volume between two or more crowders which is inaccessible to the peptides (shaded area).

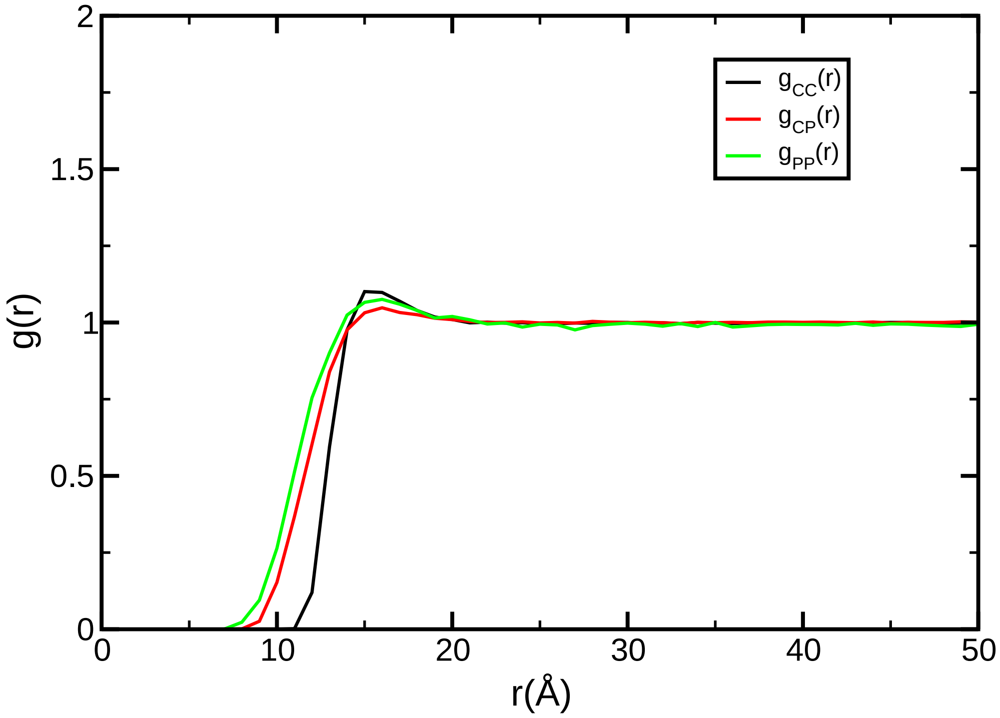


FIG. 3: Radial distribution functions of each pairs of species present in the solution (C=crowder, P=peptide).

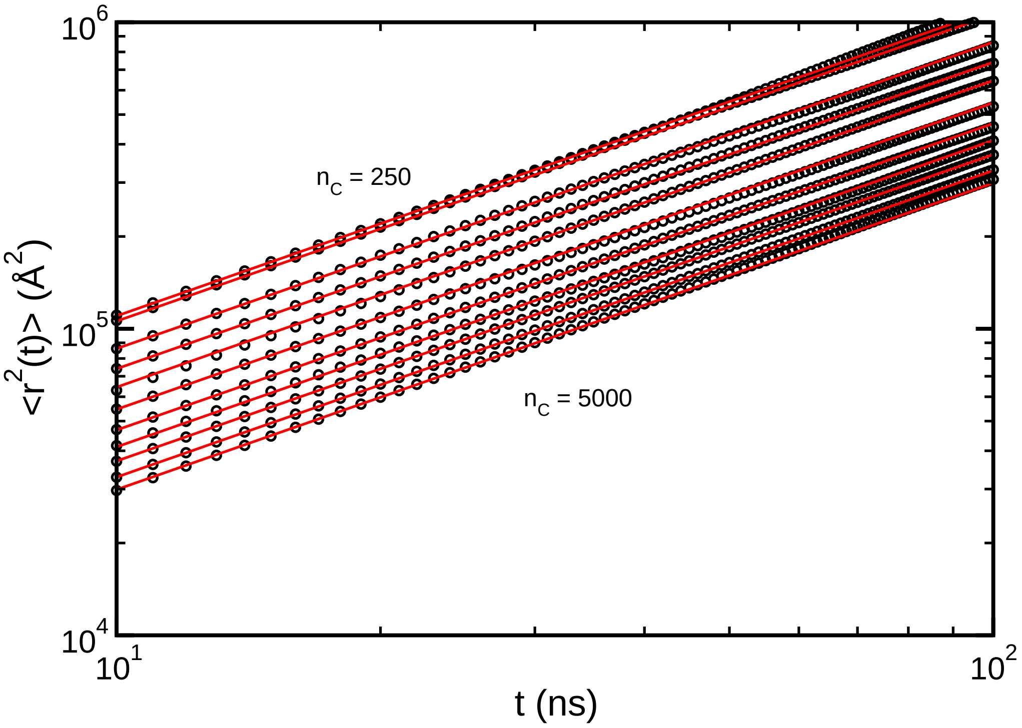


FIG. 4: Long time behaviour of the peptide mean square displacement $\langle r^2(t)\rangle$ (black circles) derived from the trajectories and linear fit with Einstein's relation (red lines) at different crowder content.

Chapter 4

Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria

Adler-Abramovich, L., Vaks, L., Carny, O., Trudler, D., Magno, A., Caflisch, A., Frenkel, D. and Gazit, E. *Nat. Chem. Biol.*, **2012**, *8, 701-706*

ARTICLE
PUBLISHED ONLINE: 17 JUNE 2012 | DOI: 10.1038/NCHEMBIO.1002

Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria

Lihi Adler-Abramovich[1], Lilach Vaks[1], Ohad Carny[1], Dorit Trudler[2,3], Andrea Magno[4], Amedeo Caflisch[4], Dan Frenkel[2,3] & Ehud Gazit[1]*

Phenylketonuria (PKU) is characterized by phenylalanine accumulation and progressive mental retardation caused by an unknown mechanism. We demonstrate that at pathological concentrations, phenylalanine self-assembles into fibrils with amyloid-like morphology and well-ordered electron diffraction. These assemblies are specifically recognized by antibodies, show cytotoxicity that can be neutralized by the antibodies and are present in the hippocampus of model mice and in parietal cortex brain tissue from individuals with PKU. This is, to our knowledge, the first demonstration that a single amino acid can form amyloid-like deposits, suggesting a new amyloidosis-like etiology for PKU.

Individuals with PKU can have mental retardation, epilepsy, organ damage and unusual posture, and the molecular mechanisms causing these are largely obscure. PKU is an autosomal recessive disorder, caused by mutations in the gene encoding phenylalanine hydroxylase (PAH) located on chromosome 12. Mutations in both alleles of the gene result in remarkably high concentrations of phenylalanine[1,2]. In untreated patients, millimolar concentrations of phenylalanine accumulate in the plasma, cerebrospinal fluid and brain tissue[3,4]. The inclusion of PKU diagnosis in newborn screening programs, by semiquantitative methods such as the Guthrie test or modern analytical tools that measure the blood concentrations of phenylalanine, allows early diagnosis of affected individuals. This permits treatment with a phenylalanine-restricted diet before the onset of clinical symptoms. A high percentage of individuals with PKU have blood phenylalanine concentrations that are above target ranges, particularly in teenagers and adults, indicating inadequate compliance to the strict diet[5]. In most of the previous studies, phenylalanine was considered to act as a neurotoxin, although the precise mechanism underlying its neurological effects still needed to be deciphered[2].

In the past decade, the role of peptide and protein aggregation in many pathological disorders was revealed. Specific attention was drawn to the formation of ordered amyloid fibrils. It was clearly demonstrated that amyloid fibrils or their early intermediates are associated with a diverse group of diseases of unrelated etiology, including Alzheimer's disease, type II diabetes and prion disorders. Despite their formation by a diverse and structurally unrelated group of proteins, all amyloid fibrils share similar biophysical and structural properties[6,7]. A variety of structural and biophysical studies indicate that aromatic residues are important in the acceleration of the amyloidogenic process and the stabilization of amyloidal structure. Although aromatic interactions are not crucial for the process of amyloid formation, they can substantially accelerate it, affect the morphology of the assemblies and reduce the minimal association concentrations[8–11]. It was previously shown that very short aromatic peptide fragments, as short as penta- and tetrapeptides, can form typical amyloid fibrils that share the same biophysical and structural properties of the assemblies formed by much larger polypeptides[9,12]. Furthermore, diphenylalanine peptide was shown to form well-ordered nanotubular assemblies by itself, with some amyloid-like structural signatures[13]. This short peptide represents the core recognition motif within the β-amyloid polypeptide, which forms amyloid plaques in Alzheimer's disease. The two phenylalanine residues (Phe19 and Phe20) in the β-amyloid peptide were suggested to mediate the intermolecular interaction between polypeptide chains. This suggestion was further substantiated by the use of phenylalanine residues as a key component of peptide-based inhibitors of β-amyloid fibril formation[14,15].

In spite of the extensive work that is performed on short peptide fragments, no study has yet examined the association between single amino acids. Inspired by the frequent occurrence of phenylalanine residues in short amyloid-forming peptides[9] and by the efficient formation of nanotubes by the self-assembly of the diphenylalanine peptide[13], we explored the ability of phenylalanine to form amyloid-like nanofibrillar structures under millimolar concentrations. As discussed below, our finding may lead to a better understanding of the pathology caused by the accumulation of phenylalanine, which can reach millimolar concentrations in individuals with PKU[1,16] and lead to new approaches of treatment for this disease.

RESULTS

Biophysical characterization of phenylalanine assemblies

We examined the ability of the aromatic amino acid phenylalanine to form ordered assemblies under pathologically relevant concentrations. We observed that phenylalanine by itself, at millimolar concentrations, self-assembles to form amyloid-like nanofibrillar structures. To study the process and outcome of phenylalanine fibril formation, we applied a series of biophysical and biological assays. Transmission electron microscopy (TEM) analysis of phenylalanine at the millimolar-concentration range indicated the occurrence of well-ordered and elongated assemblies (**Fig. 1a**). Scanning electron microscopy (SEM) was also used to study the three-dimensional structures of the fibrils (**Supplementary Results**, **Supplementary Fig. 1a**), and environmental SEM (ESEM) was used to study fibrillar structures in a humid environment (**Supplementary Fig. 1b**). Both SEM and ESEM micrographs

---

[1]Department of Molecular Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel. [2]Department of Neurobiology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel. [3]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. [4]Department of Biochemistry, University of Zurich, Zurich, Switzerland. *e-mail: ehudg@post.tau.ac.il

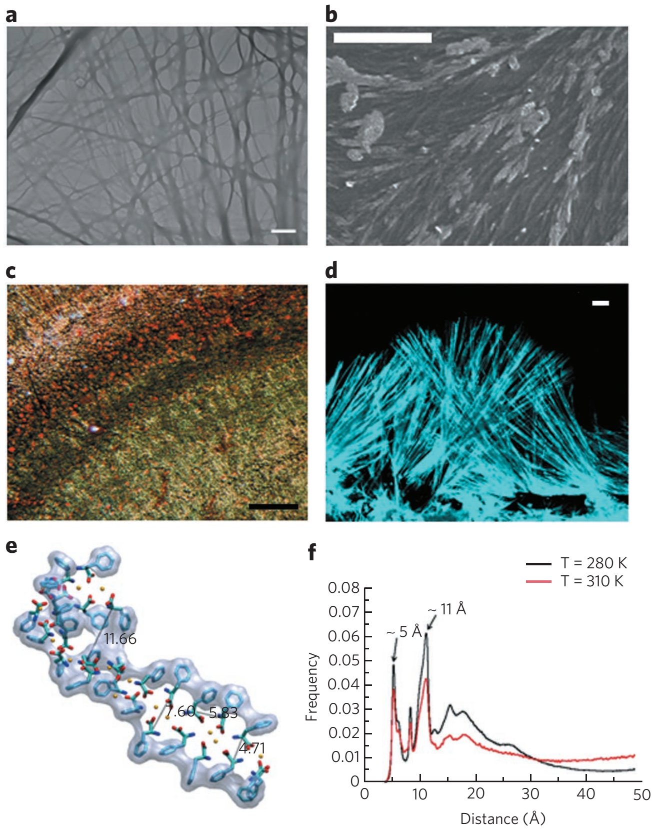


Figure 1 | The single aromatic amino acid, phenylalanine, self-assembles into supramolecular fibrillar structures. (**a**) TEM images of elongated phenylalanine fibrils. Scale bar is 1 µm. (**b**) SEM images of phenylalanine fibrils in human serum. Scale bar is 20 µm. (**c**) Microscopic examination under polarized light following Congo red staining of phenylalanine fibrils. Scale bar is 500 µm. (**d**) Confocal microscopy image of fibrils dyed with ThT. Scale bar is 10 µm. (**e**) Representative snapshot of the filamentous structure obtained by molecular dynamics simulations started with 27 monodisperse phenylalanine molecules (cyan) at high pH in the presence of counterions (yellow spheres). The tight packing of the aromatic rings is emphasized by their van der Waals envelope (gray surface). (**f**) Distribution of distances between pairs of atoms in different phenylalanine molecules in the aggregates obtained by molecular dynamics simulations. The distances between the center of mass of all pairs of the 27 phenylalanines were used for these histograms, and quantitatively similar histograms were obtained using distances between atoms instead of those between center of masses. The very similar distributions at 280 K (black) and 310 K (red) show that the ordered aggregates of phenylalanine are essentially the same in this temperature range and that the simulations have reached convergence.

showed areas covered with discrete assemblies, demonstrating that the assemblies are relatively homogeneous and are evidently discrete entities with a persistence length on the order of a few micrometers. We have also demonstrated by SEM analysis that phenylalanine fibrillar structures are formed in human serum, a more physiologically relevant environment (**Fig. 1b**). Both HPLC and NMR analysis clearly indicated that no covalent bonds were formed between the phenylalanine monomers and that the highly ordered fibrils are supramolecular assemblies (**Supplementary Fig. 1c**). All of these data imply that phenylalanine assembles into amyloid-like structures and that this assembly can take place under pathologically relevant conditions.

Another characteristic of amyloid fibrils is the presence of typical yellow-green birefringence upon staining with Congo red, visualized by microscopic examination under cross-polarized light. The birefringence results from the high order of the assemblies at the molecular level. Thus, we examined the phenylalanine assemblies using Congo red staining to gain further information on their internal order. Upon microscopic exanimation, we observed a characteristic birefringence similar to that of amyloid fibrils (**Fig. 1c**). An additional common method for quantitative assessment of amyloid fibrils is the thioflavin T (ThT) fluorescence assay, which reflects the change in the dye's fluorescence upon interaction with ordered assemblies. We used the ThT fluorescence assay to visualize the phenylalanine fibrils and observed an excitation shift typical of amyloid fibril binding of ThT. Fluorescence confocal microscopy analysis of the ThT-stained fibrils showed the presence of elongated ordered structures (**Fig. 1d**).

Another method of confirming the degree of order of fibrillar structures is the use of electron diffraction. This was previously used to probe the ultrastructure of amyloid fibrils[10,17]. Indeed, in the case of the phenylalanine fibril, electron diffraction studies also gave a strong indication to the high organization of the assemblies. An electron diffraction pattern of a single fibril was consistent with a unit cell of $a = 11.63 \pm 0.27$ Å, $c = 4.6 \pm 0.06$ Å (for $n = 5$ measurements), where a is oriented normal to the long axis of the crystal, and c is along the fiber axis (**Supplementary Fig. 1d**).

Molecular dynamics simulations[18] with a generalized Born implicit solvent model[19] were carried out to shed light on the structures of the early aggregates of phenylalanine. Multiple microsecond-long simulations were started with 27 monodispersed phenylalanine molecules at different pH values and in the presence or absence of counterions. Ordered aggregation was observed at some but not all conditions. At high pH (that is, neutral amino group and negatively charged carboxy group) in the presence of counterions, filamentous aggregates were observed in high concentrations and at all temperature values (**Fig. 1e**). Analysis of the ensemble of self-assembled structures yielded a distribution of interatomic distances with two peaks at about 5 Å and 11 Å (**Fig. 1f**), which correspond to the distance between neighboring phenylalanines and the laminal spacing, respectively (**Fig. 1e**). Pairs of neighboring phenylalanines are involved in direct hydrogen bonds or salt-bridged polar interactions. Notably, the distribution of distances is in agreement with the aforementioned electron diffraction pattern.

Taken together, the results of the set of experimental techniques and atomistic simulations further confirm that the high degree of structural order of the phenylalanine fibrillar assemblies is not the product of irregular aggregation. In this sense, they have characteristics similar to those of amyloid fibrillar deposits. We conclude that the fibrils formed by phenylalanine closely resemble amyloid structures, as confirmed by all of the physical assays used.

Cytotoxicity of phenylalanine assemblies

We examined whether, like many amyloid structures, phenylalanine assemblies have a cytotoxic effect. This was examined at a physiological range of concentrations, similar to those detected in untreated PKU individuals, by *in vitro* cellular viability experiments. To this end, elevated concentrations of phenylalanine ranging from 0.1 µM to 15 mM were added to cultured PC12 cell line (**Fig. 2a**) and to cultured Chinese hamster ovary (CHO) cells (**Supplementary Fig. 2a**). The phenylalanine fibrils, at millimolar concentrations, had a toxic effect on the cells as measured by the 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide (MTT) assay (**Fig. 2a**). In the presence of 1.8 mM phenylalanine, cell viability decreased to approximately 80%. Moreover, in the presence of 7.5 mM phenylalanine, the cells' viability was approximately 65% (**Fig. 2a**), suggesting a dose-dependent response. Our assay was performed at much shorter time scales than those for disease progression, which may account for our use of slightly elevated concentrations compared to those typically found in individuals with PKU. It is also possible

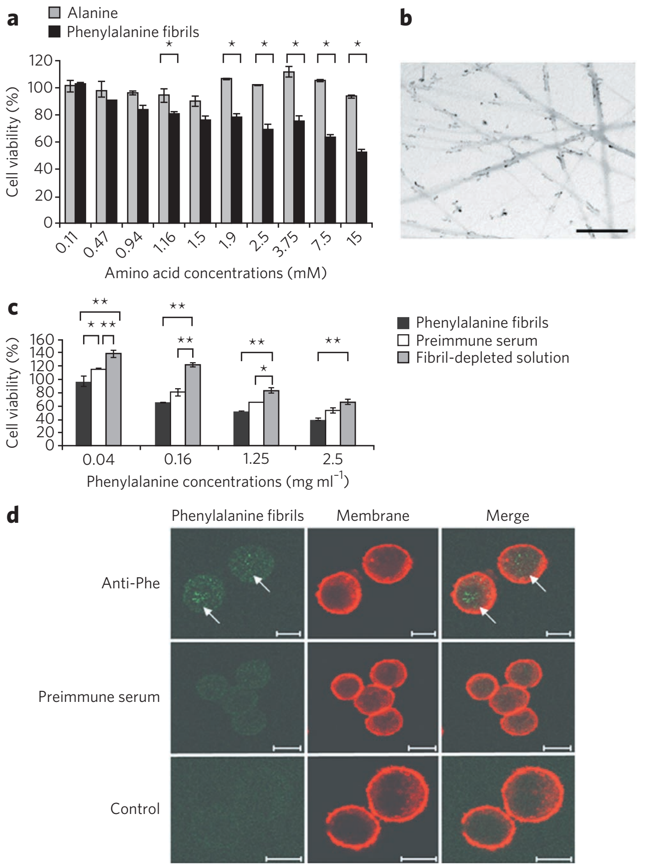


Figure 2 | Specific antibodies against phenylalanine fibrils and the toxic effect and interaction of phenylalanine fibrillar structures with cell cultures. (**a**) Cell viability was determined using the MTT assay. The PC12 cell line was maintained in the presence of phenylalanine fibrils (black bars) or the control amino acid, alanine (gray bars). $*P < 0.05$. (**b**) TEM micrographs of phenylalanine fibrils, visualized using antibodies that were specifically bound to phenylalanine fibrils and then were marked with a secondary antibody conjugated to 18-nm gold particles. Scale bar is 2 μm. (**c**) Cell viability was determined using the MTT assay, CHO cell cultures were maintained in the presence of an increasing amount of phenylalanine fibrils (black bars) or immunoprecipitated (IP) solutions of phenylalanine depleted of fibrils (gray bars). As control, phenylalanine fibrils were also incubated with preimmune serum (white bars). $*P < 0.05$; $**P < 0.001$. Results for **a** and **c** are presented as mean ± s.e.m. (**d**) CHO cell cultures were incubated with phenylalanine fibrils, fixed, and then incubated with anti-Phe and stained using Alexa 488–conjugated antibody (green, marked with arrows). The cell membrane was marked with phalloidin (red). As controls, CHO cell cultures were incubated with phenylalanine fibrils then fixed and incubated with preimmune serum or without primary antibodies. Scale bars are 10 μm.

that such high concentrations of phenylalanine occur pathologically in individuals with PKU, owing to transient local high concentrations of phenylalanine in the brain. Slightly elevated phenylalanine concentrations were used, similarly to the experimental studies of β-amyloid polypeptide, which in Alzheimer's disease occurs at overall cerebral concentrations lower than 10 nM, whereas toxicity studies are performed at the micromolar range[20,21]. Furthermore, alanine, an average-sized amino acid that does not form fibrillar structures at similar concentrations, was used as a negative control in the cytotoxicity assay and, as expected, did not demonstrate a toxic effect (**Fig. 2a**). Thus, it is quite clear that it is not merely the concentration of amino acid monomers but rather the mulitimolecular entity that is the cause of toxicity.

We further assessed the influence of the phenylalanine fibrils on the morphology of CHO cells using SEM. We observed changes in CHO cell morphology following incubation with phenylalanine fibrils. In the absence of phenylalanine fibrils, the cells had an elongated shape (**Supplementary Fig. 2b**,**c**), whereas the phenylalanine-treated cells seemed smaller and rounder (**Supplementary Fig. 2d**,**e**). We observed a very low density of the phenylalanine-treated cells compared to that of the untreated cells. On the basis of these results, we suggest that phenylalanine pathology in individuals with PKU may be due to the cytotoxic effects of fibril formation at millimolar concentrations.

Formation of antibodies against phenylalanine fibrils

One of the key prospects for amyloid disease treatment is the use of antibodies that specifically recognize and clear fibrillar assemblies and not the corresponding monomeric species. To examine whether the phenylalanine fibrils represent a unique immunological entity, we tried to produce specific antibodies against the phenylalanine fibrils. Rabbits were immunized with the fibrils, and their serum was tested for specificity to the assemblies. We employed an immunogold assay to show the antibodies' affinity to the fibrils. As can be observed in the TEM images, the fibrils were specifically marked with gold-labeled secondary antibodies bound to the serum antibodies (**Fig. 2b**). Control analysis of phenylalanine fibrils, marked with preimmune serum or only with gold-labeled secondary antibodies, did not show any specificity to the fibrils (**Supplementary Fig. 3**). Moreover, the specificity of the antibodies was demonstrated, as no cross-reactivity was found when the antibodies were incubated with the Parkinson's α-synuclein amyloid deposits or with diphenylalanine peptide nanotubes[13], and only low binding was observed when the antibodies were added to phenylalanine monomers dissolved in guanidine hydrochloride solution, reflecting their much lower ability to self-assemble. In addition, no binding was observed when preimmune serum was used as a control (**Supplementary Fig. 4**). This observation provides a clear indication for the formation of distinctive and specific immunological epitopes by the self-assembly of phenylalanine, as observed in amyloid disorders.

Immunoprecipitation depleted phenylalanine fibril toxicity

Antibodies against phenylalanine fibrils (anti-Phe fibril) were furthermore used to immunoprecipitate the phenylalanine assemblies from the solution and to assess the contribution of the assemblies to the toxic effect on the cell culture. **Figure 2c** compares the cytotoxic effect of phenylalanine assemblies prior to and following immunoprecipitation. Cells treated with solutions containing phenylalanine assemblies at concentrations of 2.5 mg ml^{-1} and 0.16 mg ml^{-1} showed only 40% and 68% cell viability, respectively; however, after immunoprecipitation, the viability of cells incubated with the solutions was elevated to 66% and 122%, respectively, as measured by the MTT assay (**Fig. 2c**). In addition, the solutions that were incubated with the cells were analyzed using electron microscopy; phenylalanine fibrils were observed in the solution before the immunoprecipitation and were not detected after immunoprecipitation with anti-Phe fibril (**Supplementary Fig. 5**). As a control, cell viability after immunoprecipitation of the phenylalanine fibrils with preimmune serum was analyzed. The viability of the cells incubated with the 2.5 mg ml^{-1} and 0.16 mg ml^{-1} solutions was moderately elevated to 59% and 80%, respectively. The nonspecific preimmune serum showed a notably lower depletion effect in comparison to the specific anti-Phe fibril depletion. The moderate elevation in viability may possibly be due to some sedimentation of the phenylalanine fibrils during the immunoprecipitation procedure. In addition, electron microscopy analysis showed phenylalanine fibrils in the preimmune immunoprecipitated solution, correlating with the results showing low viability (**Supplementary Fig. 5**).

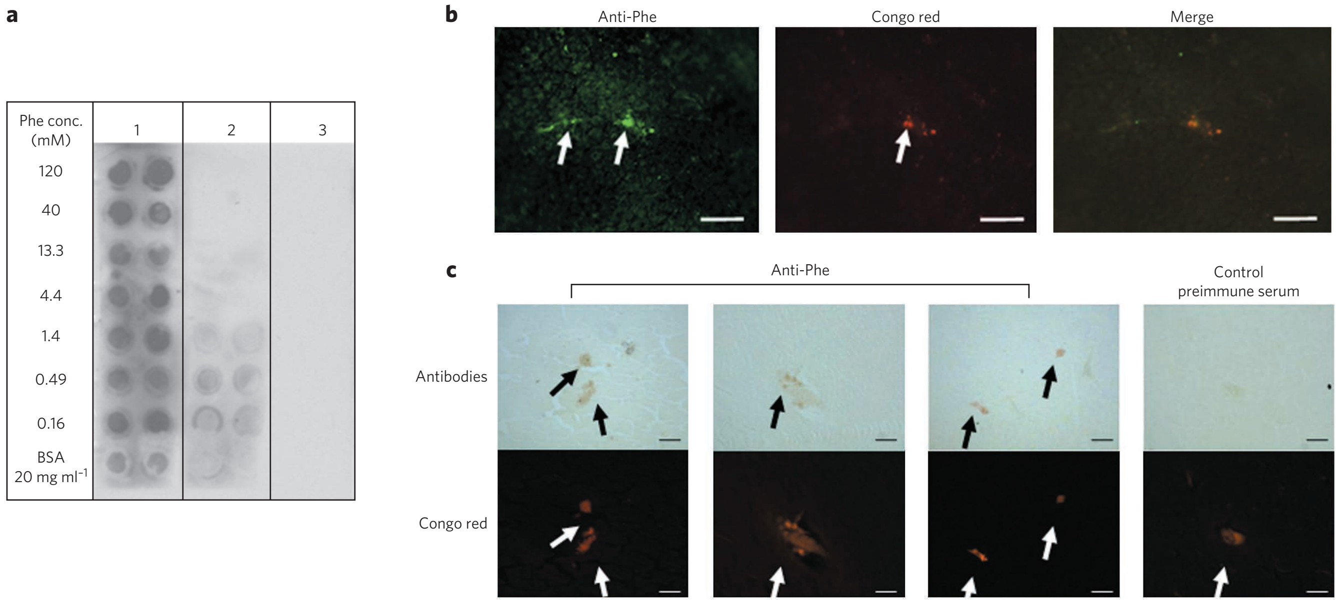


Figure 3 | Dot-blot analysis and histological staining indicates the presence of phenylalanine fibrils in model mice and PKU patient brain tissues. (**a**) Dot-blot analysis of phenylalanine fibrils. Column 1, serum of homozygous mouse (Pah^{enu2}) strongly bound the phenylalanine fibrils. Column 2, serum of heterozygous mouse (Pah^{enu2}) did not bind the phenylalanine fibrils. Column 3, serum of wild-type mouse did not bind phenylalanine fibrils. (**b**) Histological staining of homozygous Pah^{enu2} mouse brain. The 20-µm-thick brain slices were stained with rabbit anti-Phe fibril antibodies and Congo red and then were examined using fluorescent microscopy. The detected amyloid-like plaques showed colocalization of the fluorescent signal obtained from Congo red and antibody staining. (**c**) The brain of an individual with PKU was stained with anti–Phe fibril serum or with preimmune serum, examined using light microscopy and co-stained with Congo red. Phenylalanine-positive depositions were found in the parietal cortex. Scale bars are 100 µm.

Phenylalanine fibrils interact with cells

Moreover, confocal microscopy was used to study the interaction of the fibrils with CHO cells in culture (**Fig. 2d**). Cells were incubated with phenylalanine fibrils following cell fixation and staining with antibody against phenylalanine (anti-Phe) and phalloidin. Analysis of the cells clearly demonstrated that the fibrils could be detected within the cells (**Fig. 2d**) under the conditions in which cytotoxicity is observed (**Fig. 2a**). No such staining could be observed with the preimmune serum (**Fig. 2d**) or without the addition of the primary antibody (**Fig. 2d**).

The presence of phenylalanine fibrils in Pah^{enu2} mice

To prove that these assemblies exist *in vivo*, we examined serum samples obtained from a genetic mouse model of PKU (Pah^{enu2}), deficient in phenylalanine hydroxylase activity[22]. We were interested in analyzing the presence of anti-Phe fibril, which could indicate the presence of phenylalanine fibril in the PKU mouse model. Using the dot-blot assay, we were able to observe anti-Phe fibril specifically in Pah^{enu2} homozygous mouse serum (**Fig. 3a**). Moreover, anti-Phe was not detected in control Pah^{enu2} heterozygous mouse serum and in normal mouse serum (**Fig. 3a**). We further examined the presence of phenylalanine fibrils in brain tissue of Pah^{enu2} mice. We performed histology staining experiments using Congo red and Phe-specific antibody staining. The histology staining results showed evidence for the presence of amyloid-like plaques in Pah^{enu2} mice (**Fig. 3b**). A large number of plaques were specifically detected in the hippocampus and dentate gyrus (**Fig. 3b**), and in proximity to blood vessels (**Supplementary Fig. 6a**). Notably, considerable necrosis and edema in the dentate gyrus were described in previous studies on PKU[23]. In addition, the colocalization of Congo red and anti-Phe staining provides evidence that the plaques consist of amyloid-like phenylalanine fibrils. DAPI staining (nonspecific staining of double-stranded DNA) of plaque sections detected glia cell infiltration to the region (data not shown). In addition, the following four histological staining controls showed no evidence of phenylalanine assemblies**:** (i) homozygous Pah^{enu2} mouse brain tissue, stained with antibodies extracted from immunized serum depleted of its Phe fibril–specific antibodies by preincubation with phenylalanine fibrils; (ii) homozygous Pah^{enu2} mouse brain tissue, stained with antibodies derived from preimmune serum; (iii) heterozygous Pah^{enu2} mouse brain tissue stained with anti-Phe; (iv) heterozygous Pah^{enu2} mouse brain tissue stained with preimmune serum (**Supplementary Fig. 6b–e**).

Immunohistology of phenylalanine fibrils in PKU patients

We examined the presence of phenylalanine fibrils in the brain tissue of individuals with PKU. We performed histological staining with both anti-Phe fibril and Congo red. As shown in **Figure 3c**, evidence was found for the presence of phenylalanine deposition in the brain tissue using both immunostaining and Congo red staining (**Fig. 3c**). The control sample stained with Congo red showed positive staining; however, the preimmune serum did not recognize this area (**Fig. 3c**, right column). These findings demonstrate the specificity of the anti-Phe fibril. The phenylalanine assemblies' co-staining was mainly detected in the parietal cortex, which was previously suggested to be involved in the pathology of PKU in a rat model, in terms of changes in the structural organization of the cortex and decreased number of dendritic processes[24]. Thus, the phenylalanine assemblies not only are clear supramolecular entities but also are most relevant to the disease, as determined by the mouse model experiments.

DISCUSSION

The current study presents a new paradigm to explain the pathology of PKU and suggest new routes for potential therapy. Our study indicates, for what is to our knowledge the first time, the ability of phenylalanine, a single amino acid, to form well-ordered fibrillar assemblies at the nano scale. These assemblies are not irregular aggregates, as they have typical fibrillar morphology, characteristic birefringence, ThT fluorescence patterns and, above all, clear electron diffraction patterns. In all of these aspects, the properties of these fibrils highly resemble those of amyloid assemblies, which are related to numerous pathological disorders. The formed structures not only are ordered as amyloid fibrils but also have strong and clear cytotoxic activities, as other amyloid assemblies do[25,26]. Moreover, the formation of phenylalanine aggregates could be detected in the brain of PKU model mice and individuals with PKU using Phe-specific antibodies. These findings suggest that PKU is closely related to the family of amyloid-related diseases and might have similar etiology. Moreover, the phenylalanine fibrils represent a distinct immunological entity, like amyloid assemblies, and many concepts and experimental studies that are used for the development of immunological treatment for amyloid diseases may also be used in this case[27,28].

METHODS

Materials. Amino acids were purchase from Sigma (purity ≥98%). Fresh stock solutions were prepared by dissolving the amino acid in double-distilled H_2O (ddH_2O), PBS or DMEM (Beit Haemek) or human serum at various concentrations ranging from 6 μM to 120 mM.

TEM. Phenylalanine was dissolved in ddH_2O to a concentration of 6 mM and incubated at 25 °C for 2 h. A 10-μl aliquot of this solution was placed on 400-mesh copper grids. After 1 min, excess fluids were removed. For negative staining, the grid was stained with 2% (w/v) uranyl acetate in water, and after 2 min excess fluids were removed from the grid. Samples were viewed using a JEOL 1200EX electron microscope operating at 80 kV.

Molecular dynamics simulations. Molecular dynamics simulations with a generalized Born implicit solvent model[19] were carried out. Multiple microsecond-long simulations were started with 27 monodispersed phenylalanine molecules at different pH values and in the presence or absence of counterions. Four different concentrations of phenylalanine (1 mM, 6 mM, 30 mM and 100 mM) and three temperature values (280 K, 300 K and 310 K) were used.

Antibodies formation. Phenylalanine was dissolved in ddH_2O at a concentration of 120 mM to form fibrils; rabbits were immunized five times subcutaneously at 14-d intervals with the fibrils and Freund's adjuvants. Seven days after each injection, the rabbits were bled, and their serum was tested using dot-blot analysis.

Rabbit antibody immunotesting using TEM analysis. The immunolabeling was visualized using goat anti-rabbit conjugated with 18-nm gold (Jackson ImmunoResearch; catalog no. 111-215-144 1:20). Phenylalanine was adhered to the copper grid as described in the TEM section. Then, the grid was blocked with 1% (w/v) BSA and 3% (w/v) goat serum for 30 min. Samples were incubated with the serum diluted 1:200 in PBS in 1% milk for 30 min, washed five times with 0.1% BSA, then incubated with the secondary antibody for 30 min and similarly washed. Samples were viewed using a JEOL 1200EX electron microscope operating at 80 kV.

Immunoprecipitation and cell cytotoxicity experiments. Phenylalanine solutions at concentrations of 2.5 mg ml^{-1}, 1.25 mg ml^{-1}, 0.16 mg ml^{-1} and 0.04 mg ml^{-1} were either incubated with CHO cells or immunoprecipitate, and then the solutions, with or without the fibrils, were incubated with CHO cells according to the cell cytotoxicity experiments detailed in the **Supplementary Methods**. Samples were immunoprecipitated with either anti-Phe fibril or preimmune serum (1:10) overnight at 4 °C (previously purified on protein A column, in PBS plus 2% (w/v) BSA).

Interaction of phenylalanine fibrils with CHO cells. The interaction of the fibrils with cells in culture was visualized by immunocytochemistry staining. Cells were seeded on a cover slip in a 24-well plate. Cells were washed with PBS and fixed with 4% (v/v) paraformaldehyde (PFA) in PBS for 30 min at 25 °C and then were washed twice with PBS and permeabilized with 0.1% Triton in PBS for 2 min. Following two PBS washes, cells were blocked with 10% (v/v) normal goat serum in 3% (**w/v**) BSA for 30 min and incubated with anti-Phe diluted 1:1,000 and 4 μg ml^{-1} phalloidin (Sigma) for 1 h, followed by an additional hour of incubation with Alexa 488–conjugated goat anti–rabbit IgG (Jackson ImmunoResearch; catalog no: 711-545-152; diluted 1:500). After being thoroughly washed with PBS, cells were mounted using ProLong Antifade (Invitrogen). Images were taken with LSM META confocal microscope (Zeiss).

Congo red staining of brain tissue of Pah^{enu2} mice. Fifteen-micrometer coronal brain sections of Pah^{enu2} homozygous and heterozygous mice were prepared using cryostat. Brain samples were fixed in 70% (v/v) ethanol for 1 min, washed in ddH_2O for 2 min and stained with previously filtered Congo red solution for 10 min. Following staining, the samples were washed in ddH_2O for 2 min and washed 8–10 times in NaOH-ethanol solution (0.5 ml 1% (w/v) NaOH plus 49.5 ml 50% (v/v) ethanol) until the excess red color disappeared. Finally, the samples were washed in ddH_2O, and the signal was detected using fluorescent microscopy (absorption at 498 nm, emission at 614 nm).

Immunohistological staining of mouse brain tissue. Black and tan, brachyuric (BTBR)-Pah^{enu2} brain samples were fixed in 4% PFA (in PBS) for 5 min and washed for 5 min in ddH_2O. The slices were blocked with 2% BSA solution (in PBS) for 20 min, then washed three times in ddH_2O and incubated with rabbit anti-Phe fibril solution, dilution 1:20 (previously purified on a protein A column, in PBS plus 2% (w/v) BSA) for 1 h at 25 °C. Following the incubation, the samples were washed three times in ddH_2O and incubated with goat rabbit-specific secondary antibody conjugated to Alexa F488 (Invitrogen; catalog no. A11008; diluted 1:250). Signal was detected by fluorescent microscope (absorption at 495 nm, emission at 519 nm).

Immunohistological staining of human brain tissue. Brain samples, acquired from the London Neurodegenerative Diseases Brain Bank (part of BrainNet Europe), were fixed in paraffin. Consecutive sections were deparaffinized with xylene, fixed in 4% (v/v) PFA and treated with 0.3% (v/v) H_2O_2 (in PBS). Sections were then heated in citric acid (pH 6) for 5 min and were treated with 0.25% (v/v) Triton X-100 for 3 min. The sections were blocked using 2% (w/v) BSA solution (in PBS) for 20 min, then washed three times in PBS and incubated with rabbit anti-Phe fibril solution, dilution 1:50 (previously purified on protein A column, in PBS plus 2% (w/v) BSA) for 1 h at 25 °C, along with preimmune rabbit serum as control. Following the incubation, the samples were washed 3 times in PBS and incubated with goat rabbit-specific secondary antibody conjugated to biotin (Vector laboratories, BA-1000) diluted 1:250 for 1 h at 25 °C and washed with PBS. Sections were then treated with ABC reagent (Vector Laboratories, Vectastatin ABC kit, PK-6100) and developed with diaminobenzidine and hydrogen peroxide (Vector Laboratories, SK-4100). Signal was detected by light microscope.

Statistical analysis. Two-tailed Student's t-test was performed when two groups were compared. The one-way analysis of variance followed by Bonferroni's multiple comparison tests was carried out for multiple samples. Statistical significance was determined at $P < 0.05$.

Received 26 July 2010; accepted 7 May 2012;
published online 17 June 2012

References

1. Hanley, W.B. Adult phenylketonuria. *Am. J. Med.* **117**, 590–595 (2004).
2. Surtees, R. & Blau, N. The neurochemistry of phenylketonuria. *Eur. J. Pediatr.* **159** (suppl.), S109–S113 (2000).
3. Choi, T.B. & Pardridge, W.M. Phenylalanine transport at the human blood-brain barrier. Studies with isolated human brain capillaries. *J. Biol. Chem.* **261**, 6536–6541 (1986).
4. Krause, W. *et al.* Biochemical and neuropsychological effects of elevated plasma phenylalanine in patients with treated phenylketonuria. A model for the study of phenylalanine and brain function in man. *J. Clin. Invest.* **75**, 40–48 (1985).
5. MacDonald, A., Gokmen-Ozel, H., van Rijn, M. & Burgard, P. The reality of dietary compliance in the management of phenylketonuria. *J. Inherit. Metab. Dis.* **33**, 665–670 (2010).
6. Chiti, F. & Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
7. Rochet, J.C. & Lansbury, P.T. Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* **10**, 60–68 (2000).
8. Inouye, H., Sharma, D., Goux, W.J. & Kirschner, D.A. Structure of core domain of fibril-forming PHF/Tau fragments. *Biophys. J.* **90**, 1774–1789 (2006).
9. Gazit, E. A possible role for π-stacking in the self-assembly of amyloid fibrils. *FASEB J.* **16**, 77–83 (2002).
10. Makin, O.S. & Serpell, L.C. Structures for amyloid fibrils. *FEBS J.* **272**, 5950–5961 (2005).
11. Gazit, E. Global analysis of tandem aromatic octapeptide repeats: the significance of the aromatic-glycine motif. *Bioinformatics* **18**, 880–883 (2002).
12. Reches, M., Porat, Y. & Gazit, E. Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *J. Biol. Chem.* **277**, 35475–35480 (2002).
13. Reches, M. & Gazit, E. Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **300**, 625–627 (2003).
14. Tjernberg, L.O. *et al.* Arrest of β-amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.* **271**, 8545–8548 (1996).
15. Soto, C. β-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy. *Nat. Med.* **4**, 822–826 (1998).

16. Hörster, F. *et al.* Phenylalanine reduces synaptic density in mixed cortical cultures from mice. *Pediatr. Res.* **59**, 544–548 (2006).
17. Makin, O.S., Atkins, E., Sikorski, P., Johansson, J. & Serpell, L.C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. USA* **102**, 315–320 (2005).
18. Brooks, B.R. *et al.* CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
19. Haberthür, U. & Caflisch, A. FACTS: fast analytical continuum treatment of solvation. *J. Comput. Chem.* **29**, 701–715 (2008).
20. Lesné, S. *et al.* A specific amyloid-β protein assembly in the brain impairs memory. *Nature* **440**, 352–357 (2006).
21. Abramov, E. *et al.* Amyloid-β as a positive endogenous regulator of release probability at hippocampal synapses. *Nat. Neurosci.* **12**, 1567–1576 (2009).
22. Shedlovsky, A., McDonald, J.D., Symula, D. & Dove, W.F. Mouse models of human phenylketonuria. *Genetics* **134**, 1205–1210 (1993).
23. Gazit, V., Ben-Abraham, R., Pick, C.G. & Katz, Y. β-Phenylpyruvate induces long-term neurobehavioral damage and brain necrosis in neonatal mice. *Behav. Brain Res.* **143**, 1–5 (2003).
24. Cordero, M.E., Trejo, M., Colombo, M. & Aranda, V. Histological maturation of the neocortex in phenylketonuric rats. *Early Hum. Dev.* **8**, 157–173 (1983).
25. Lashuel, H.A., Hartley, D., Petre, B.M., Walz, T. & Lansbury, P.T. Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* **418**, 291 (2002).
26. Bucciantini, M. *et al.* Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507–511 (2002).
27. Solomon, B. Clinical immunologic approaches for the treatment of Alzheimer's disease. *Expert Opin. Investig. Drugs* **16**, 819–828 (2007).
28. Schenk, D. *et al.* Immunization with amyloid-β attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* **400**, 173–177 (1999).

Acknowledgments

We thank R. Shaltiel-Karyo for confocal microscopy analysis, S. Wolf for the electron diffraction analysis, L. Buzhansky for help with the NMR and HPLC analysis, J. Delarea for help with TEM and SEM experiments, Z. Barkay for help with the SEM and ESEM analysis, S.-C. Jung (Ewha Womans University, Korea) for BTBR-*Pah*^{enu2} mouse plasma and tissue samples, C. Troakes (London Neurodegenerative Diseases Brain Bank, King's College London and part of BrainNet Europe) and T. Arzberger (Centre for Neuropathology and Prion Research, München) for brain tissue samples, I. Benhar and members of the Gazit laboratory for helpful discussions. L.A.-A. gratefully acknowledges the support of the Colton Foundation. This work was partly supported by the Israel Science Foundation–Legacy Heritage Biomedical Science Partnership grant 862/09 and the Alzheimer's Association grant NIRG-11-205535 (to D.F.). The work in the A.C. group was supported by the Swiss National Science Foundation.

Author contributions

L.A.-A., O.C. and E.G. conceived and designed the experiments. L.A.-A. and L.V. planned and performed the experiments. D.T., L.V. and L.A.-A. designed and performed the mouse and human histology experiments. D.F. designed and coordinated the mice and human histology experiments. A.M. designed and performed the molecular dynamics simulations. A.C. designed and coordinated the molecular dynamics simulations. L.A.-A. and E.G. wrote the paper. All authors discussed the results and commented on the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to E.G.

Supplementary Information

Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria

Lihi Adler-Abramovich[±], Lilach Vaks[±], Ohad Carny[±], Dorit Trudler[#], Andrea Magno[¶], Amedeo Caflisch[¶], Dan Frenkel[#] and Ehud Gazit[±]*

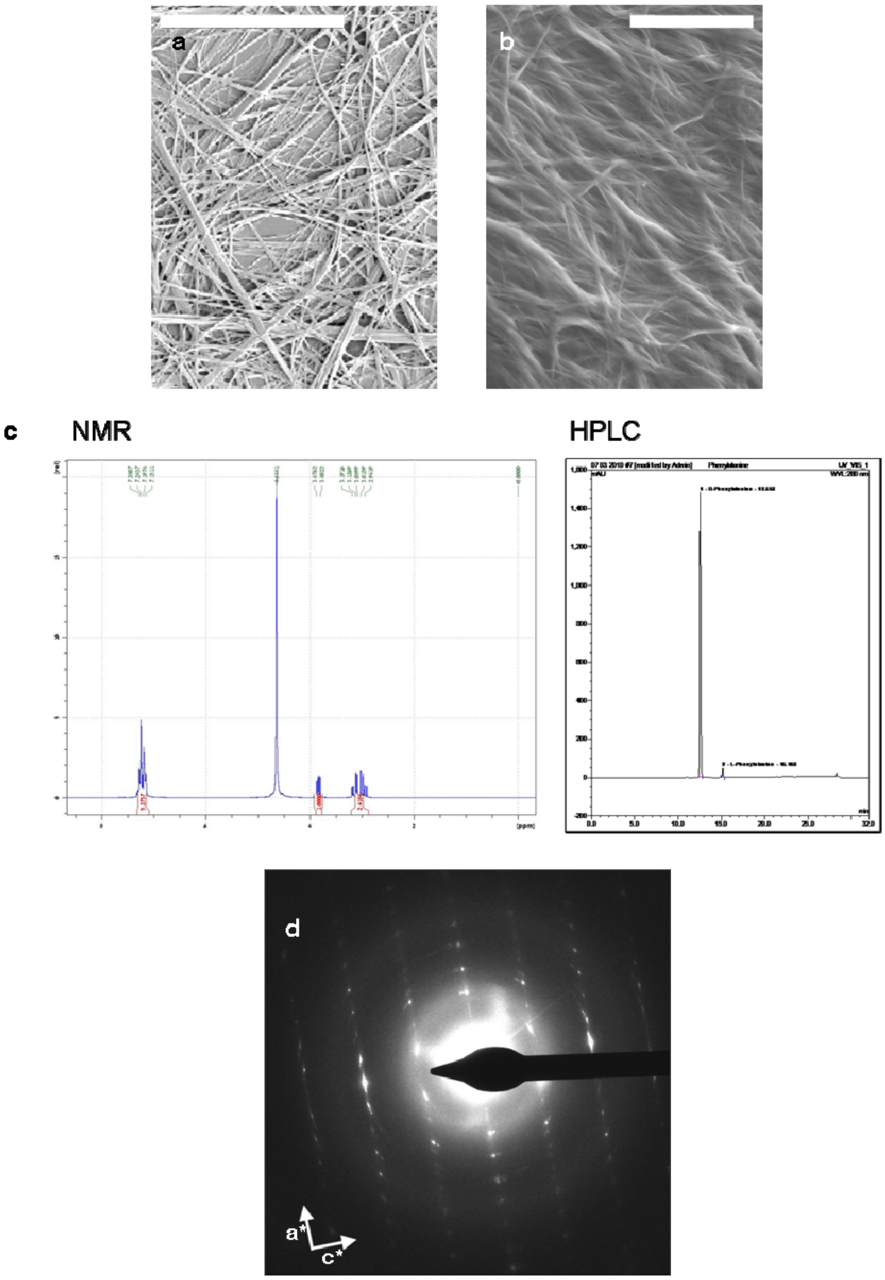
[±]Department of Molecular Microbiology and Biotechnology, [#]Department of Neurobiology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel, [¶]Department of Biochemistry, University of Zurich, Zurich, Switzerland

[±]Department of Molecular Microbiology and Biotechnology, [#]Department of Neurobiology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv 69978, Israel.

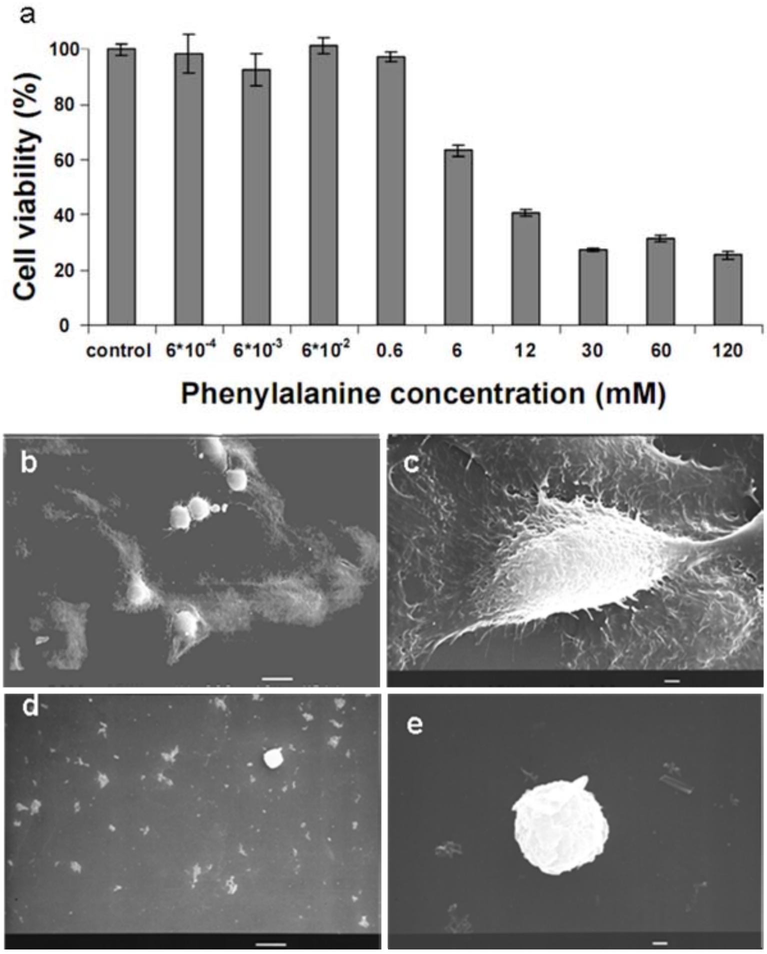
Supplementary Results

Monomer concentration at equilibrium:

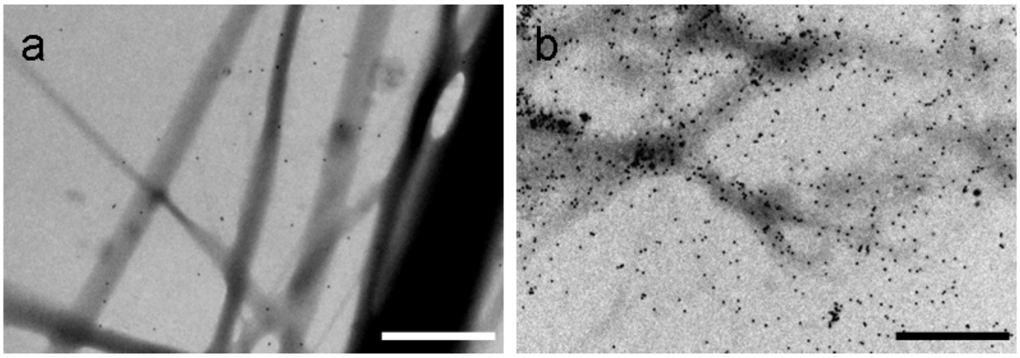
We measured the phenylalanine absorption at 256 nm and calculated its concentration. The absorbance of the assemblies' solution was 0.856 and the calculated concentration was 5 mM. Then, the assemblies were sediment using ultra-centrifugation and the supernatant monomers solution absorbance at 256 nm was 0.08. Therefore, the calculated concentration was 0.476 mM.



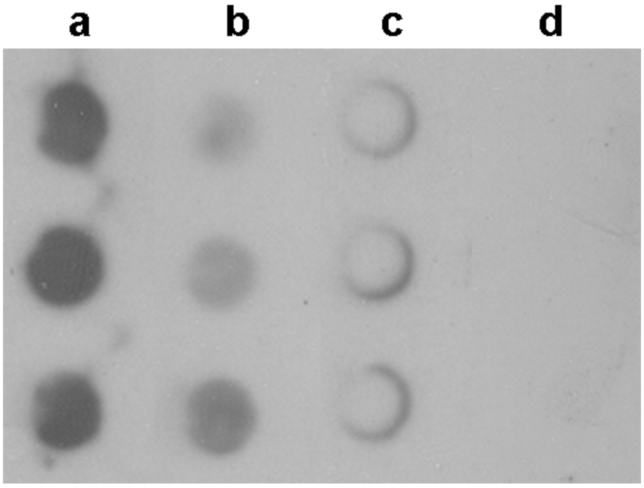
Supplementary Figure 1: (**a**) Scanning electron microscopy (SEM) image of the phenylalanine fibrils. (**b**) Environmental scanning electron microscopy image of the phenylalanine fibrils. Scale bars for a, b are 20 µm. (**c**) HPLC and NMR analysis of phenylalanine assemblies indicating that no covalent bonds between the phenylalanine monomers were formed. (**d**) Electron-diffraction analysis of a single fibril. Axis a* is oriented normal to the long axis of the crystal. a* and c* are the reciprocal lattice vectors in the diffraction pattern.



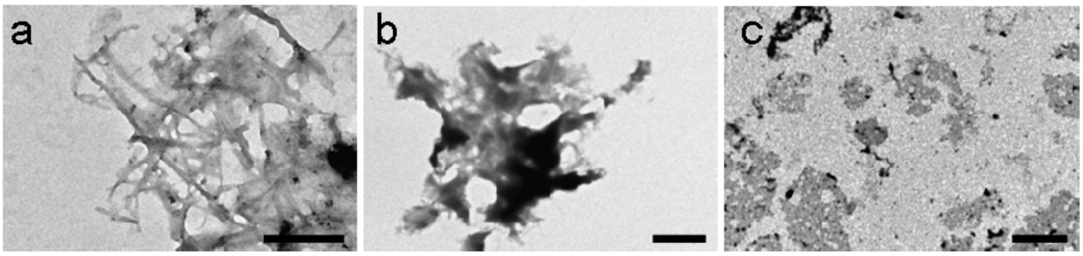
Supplementary Figure 2: The toxic effect of phenylalanine fibrillar structures on CHO cells. (**a**) Cell viability was determined using the MTT assay. CHO cell line was maintained in the absence or presence of increasing amounts of phenylalanine fibrils. (**b-c**) Scanning electron microscopy (SEM) images of untreated CHO cells. (**d-e**) SEM image of CHO cells incubated with phenylalanine fibrils. Scale bars for b, d are 10 μm and for c, e are 1 μm.



Supplementary Figure 3: Transmission electron microscopy of phenylalanine fibrils with the presence of secondary antibodies conjugated to 18 nm gold particles. (**a**) In the absence of primary antibodies. This control sample does not show specific binding to the fibril. (**b**) In the presence of pre-immune serum. The non-specific antibodies were not bound to the phenylalanine fibrils. Scale bars are 1 μm.

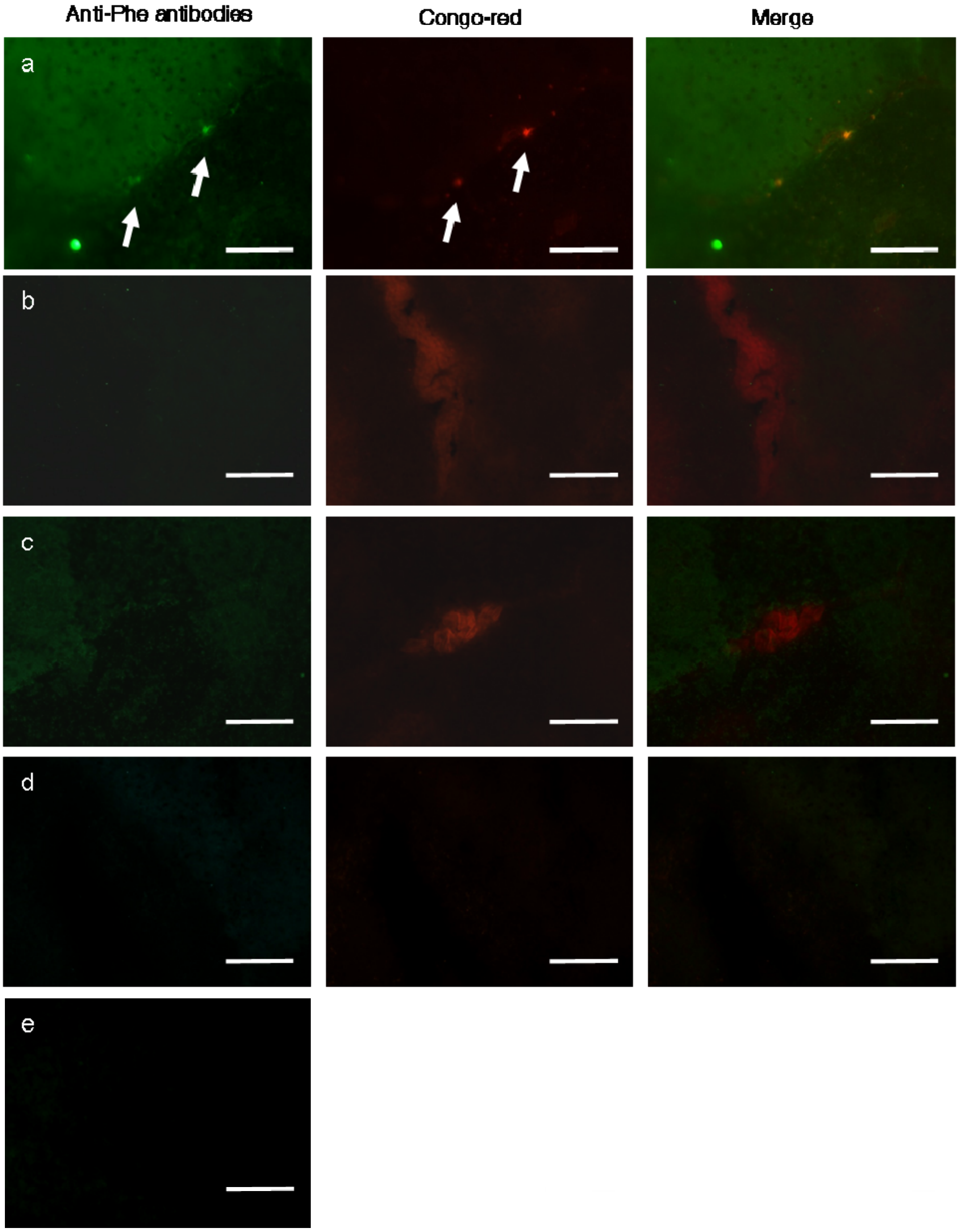


Supplementary Figure 4: Dot-blot binding analysis of the immune and pre-immune serum. (**a**) Serum from immunized rabbit was bound to the phenylalanine fibrils and demonstrated the specific binding of the serum antibodies to the phenylalanine fibrils. (**b**) Serum from immunized rabbit was incubated with phenylalanine monomers dissolved in guanidine hydrochloride solution reflecting a much lower self-assembly ability. (**c**) Serum from immunized rabbit was incubated with diphenylalanine peptide nanotubes, and demonstrated the inability of the antibodies to bind to the diphenylalanine peptide nanotubes. (**d**) Pre-immune serum was incubated with phenylalanine fibrils, and demonstrated the inability of the antibodies to bind to the fibrils. All the rows represents decreasing concentrations from top to bottom.



Supplementary Figure 5: Transmission electron microscopy micrographs of the solution that were incubated with CHO cells (**a**) Phenylalanine fibrils were observed prior to the IP. (**b**) Phenylalanine fibrils were observed following the IP with pre-immune serum. (**c**) Phenylalanine fibrils were not detected after IP with anti-Phe fibril antibodies. Scale bars for a, c are 500 nm and for b is 2 μm.

Nature Chemical Biology: doi:10.1038/nchembio.1002



Supplementary Figure 6: Histological staining of homozygous and heterozygous pah^{enu2} mice brains. **(a)** Homozygous *Pah*^{enu2} mouse brain slices were stained with rabbit anti–Phe fibril antibodies and Congo red and then were examined using fluorescent microscopy. The detected amyloid-like plaques showed colocalization of the fluorescent signal obtained from Congo red and antibody staining in proximity to blood vessels. (**b**) Homozygous mouse brain slices were stained with immunized serum that was antibody depleted (**c**) Homozygous mouse brain slices were stained with antibodies from pre-immune serum (**d**) Heterozygous mouse brain stained with immunized serum (**e**) Heterozygous mouse brain stained with pre-immune serum. b-d were stained with Congo red, then examined using fluorescent microscopy. b-e Do not show any specific staining. Scale bars are 100 µm.

Supplementary Methods

Scanning electron microscopy. Phenylalanine was dissolved in ddH$_2$O or 10% human serum to a concentration of 6 mM or 12mM respectively, and incubated at room temperature for two hours. A 10 µl aliquot of the solution was placed on microscope glass cover slip and coated with gold. Scanning electron microscopy images were taken using a JSM JEOL 6300 SEM operating at 5 kV.

Environmental scanning electron microscopy. Phenylalanine was dissolved in ddH$_2$O to a concentration of 6 mM and incubated at room temperature for two hours. A 10 µl aliquot of the solution was placed on a metal stand. Environmental scanning electron microscopy images were taken using Quanta 200 FEG Field Emission Gun ESEM operating at 10 kV.

Congo red staining and birefringence. Phenylalanine was dissolved in ddH$_2$O to a concentration of 6 mM and incubated at room temperature for two hours. A 10µl aliquot of the solution was allowed to dry on a glass microscope slide. Staining was performed by the addition of 10 µl solution of 80% ethanol saturated with Congo red and NaCl. Birefringence was determined with a SZX-12 Stereoscope (Olympus, Hamburg, Germany) equipped with a polarizing stage.

ThT staining and confocal laser microscopy imaging. 10 µl ThT solution (2 mM, PBS buffer) were mixed with 10 µl phenylalanine fibril (6 mM, ddH$_2$O). An LSM 510 confocal laser scanning microscope (Carl Zeiss Jena, Germany) was used at excitation and emission wavelengths of 458 and 485 nm, respectively.

Electron diffraction. Phenylalanine was dissolved in ddH_2O to a concentration of 6 mM and incubated at room temperature for two hours. Then, a 10 µl aliquot of this solution was placed on 400 mesh copper grids. After 1 minute, excess fluid was removed. Electron diffraction experiments were performed on an FEI Tecnai F20 microscope FEI at 200 kV with a field-emission gun, and samples cooled to liquid nitrogen temperatures using a Gatan 626 cryoholder. Low-dose methods were used with total dose to the sample of ~50 electrons per $Å^2$. Electron diffraction patterns were recorded directly to the CCD camera (TVIPS F415).

NMR. NMR spectra were recorded on a AC 200MHz, Bruker spectrometer, using Bruker Topspin 2.1 software. The chemical shifts were expressed in δ relative to TMS (δ=0 ppm). The spectra were recorded in D_2O as a solvent, at room temp. 1H-NMR (D_2O-*d6*): δ = 2.9-3.3 (m, CH2), 3.8-3.9 (m, CH), 7.1-7.3 (m,5H aromatic).

HPLC. Reverse phase HPLC showed >97% purity. Dionex HPLC system with Ultimate 3000 pump, Ultimate 3000 autosampler and Ultimate 3000 variable multiwave detector, controlled via Chromeleon chromatography workstation. Column: LiCroCART Purospher STAR RP 4.6 mm, 5 µm C18e, Buffer A: 0.1% TFA in water, Buffer B: 0.1% TFA in acetonitrile, Flow: 1 ml/min, Binary Gradient: t=0-5; %A=100%, t=5-20; %B=0%-100%, t=20-25, %B=100%, 260 nm, Diluent 100% A, TFA (Sigma) ≥99.0% (GC), for HPLC, Acetonitrile (Bio-Lab) HPLC grade, Water (Bio-Lab) HPLC grade.

Phenylalanine concentration at equilibrium. Phenylalanine was dissolved in water to a concentration of 1 mg/ml, the solution absorbance was measured at 256 nm. Then the assemblies were centrifuged in Optima TLXl Benchtop Ultracentrifuge for 1 hour at 4°C. The supernatant absorbance was measured at 256 nm.

Cell cytotoxicity experiments. PC12 and CHO cells (2×10^5 cells/mL) were cultured in 96-well micro plates (100 µL/well) and incubated overnight at 37°C. 100 µL of phenylalanine dissolved in Dulbecco's Modified Eagle Medium (DMEM) (Beit Haemek, Israel) at various concentrations were added to each well. Each experiment was repeated 3 times. Following incubation for 6 hours at 37°C, cell viability was evaluated using the 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide (MTT) assay. Briefly, 20 µL of 5 mg/mL MTT dissolved in PBS was added to each well. After 4 hours incubation at 37°C, 100 µL of the extraction buffer [20% SDS dissolved in a solution of 50% dimethylformamide and 50% DDW (pH 4.7)] were added to each well, and the plates were incubated again overnight at 37°C. Finally, color intensity was measured using an ELISA reader at 570 nm. Results are presented as mean +/- the standard error of the mean.

Scanning electron microscopy imaging of cells. CHO cell were cultured on glass cover slips located in 24-well micro plates, then incubated with various concentration of phenylalanine fibrils for 6 hour at 37 ºC, as described in the cell cytotoxicity experiments section. Cells were then fixated on the glass cover slip with 2.5% glutaraldehyde, then dehydrated and coated with gold. Scanning electron microscopy images were made using a JSM JEOL 840A SEM operating at 5 kV.

Rabbits Antibodies' Immuno-testing using dot-blot analysis. 100 µl solution of phenylalanine fibrils (0.125, 0.062, 0.32 mg/ml) were applied via a vacuum manifold onto a nitrocellulose filter using a dot-blot apparatus. After blocking the membrane with 1% milk in TBS for 1 hour at room temperature, the membrane was briefly washed with TBS and incubated with the serum diluted 1: 500 in TBS / 2% milk for 1 hour at room temperature. Then, the membrane was briefly washed with TBS and incubated with HRP-conjugated goat anti rabbit antibodies. The membrane was developed using ECL reagents (NEN, USA) according to the supplier's instructions. The blot was visualized using X-ray film. For the control binding analysis of rabbit's antibodies (1) diphenylalanine peptide (Bachem) was dissolved in water at 65 °C for 30 minutes, then cooled at room temperature to form the peptide nanotubes. (2) Phenylalanine was dissolved in 6M guanidine hydrochloride. (3) Phenylalanine fibrils that were applied to nitrocellulose were incubated with pre-immune serum. The solutions were applied via a vacuum manifold onto a nitrocellulose filter using a dot-blot apparatus followed by the protocol stated above.

Dot blot analysis of pah^{enu2} mice. Proteins or samples in a total volume of 100 µl ddH_2O or PBS were applied via a vacuum manifold onto a PVDF membrane filter using a dot–blot apparatus (Schleicher and Schuell, USA). After blocking the membranes with 3% (v/v) non-fat milk in PBS for overnight at 4ºC, the membrane was briefly washed with PBS followed by incubation with mouse (pah^{enu2} homozygous, pah^{enu2} heterozygous and wild-type) plasma or specific anti-Phe fibrils rabbit serum (previously purified on protein A column, in PBS + 2% BSA) for 1h at room temperature. The signal was detected using the appropriate anti-mouse and anti-rabbit HRP-conjugated secondary antibodies (Jackson Laboratories, West Grove, PA). After three washes with PBS the PVDF filter membranes were developed with the ECL reagent (Pierce, USA). The protein blot was visualized using X-ray film.

Chapter 5

Carnosine inhibits $A\beta_{42}$ aggregation by perturbing the H-bond network in and around the central hydrophobic cluster

Attanasio, F., Convertino, M., Magno, A., Caflisch, A., Corazza, A., Haridas, H., Esposito, G., Cataldo, S., Pignataro, B., Milardi, D. and Rizzarelli, E. *ChemBioChem*, **2013**, *14, 583-592*

DOI: 10.1002/cbic.201200704

Carnosine Inhibits $A\beta_{42}$ Aggregation by Perturbing the H-Bond Network in and around the Central Hydrophobic Cluster

Francesco Attanasio,[a] Marino Convertino,[b] Andrea Magno,[b] Amedeo Caflisch,[b] Alessandra Corazza,[c] Haritha Haridas,[c] Gennaro Esposito,[c] Sebastiano Cataldo,[d] Bruno Pignataro,[d] Danilo Milardi,*[a] and Enrico Rizzarelli[a, e]

Aggregation of the amyloid-β peptide (Aβ) into fibrillar structures is a hallmark of Alzheimer's disease. Thus, preventing self-assembly of the Aβ peptide is an attractive therapeutic strategy. Here, we used experimental techniques and atomistic simulations to investigate the influence of carnosine, a dipeptide naturally occurring in the brain, on Aβ aggregation. Scanning force microscopy, circular dichroism and thioflavin T fluorescence experiments showed that carnosine does not modify the conformational features of Aβ42 but nonetheless inhibits amyloid growth. Molecular dynamics (MD) simulations indicated that carnosine interacts transiently with monomeric Aβ42 by salt bridges with charged side chains, and van der Waals contacts with residues in and around the central hydrophobic cluster (^{17}LVFFA^{21}). NMR experiments on the nonaggregative fragment Aβ12–28 did not evidence specific intermolecular interactions between the peptide and carnosine, in agreement with MD simulations. However, a close inspection of the spectra revealed that carnosine interferes with the local propensity of the peptide to form backbone hydrogen bonds close to the central hydrophobic cluster (residues E22, S26 and N27). Finally, MD simulations of aggregation-prone Aβ heptapeptide segments show that carnosine reduces the propensity to form intermolecular backbone hydrogen bonds in the region 18–24. Taken together, the experimental and simulation results (cumulative MD sampling of 0.2 ms) suggest that, despite the inability of carnosine to form stable contacts with Aβ, it might block the pathway toward toxic aggregates by perturbing the hydrogen bond network near residues with key roles in fibrillogenesis.

Introduction

Alzheimer's disease (AD), the most common form of senile dementia, is a neurodegenerative disease characterized by the presence of neurofibrillary tangles and amyloid plaques in the brain.[1] The main constituents of the plaques are two peptides (40 and 42 amino acids long) called amyloid β peptides (Aβ40 and Aβ42).[2] At first, attention was entirely focused on amyloid fibrils as the cause of AD, but more recent studies have suggested that the small-sized, soluble Aβ oligomers formed during the early steps of peptide aggregation are the main cytotoxic agents,[3] whereas Aβ monomers show a neuroprotective role.[4] Although a causal relationship between the morphology of Aβ aggregates and the severity of the disease remains open to discussion, considerable evidence indicates that a key event in AD pathogenesis is the conversion of Aβ from its neuroprotective, soluble, monomeric state[4] into various toxic aggregated assemblies in the brain. Therefore, the design of effective molecules to prevent the self-assembly of Aβ into toxic oligomers is widely recognized as the primary goal of a number of therapeutic strategies under development or in clinical trials. Despite the identification of oligomers as the critical pathogenic form of Aβ aggregates, uncertainty concerning the structural features triggering toxic oligomerization persists due, in part, to the transient lifetimes of early intermediates and the dynamic transition into the thermodynamically favored fibrils. Recently, Ahmed et al. reported an Aβ42 fibril model with a turn spanning residues 24–27 in both oligomers and fibrils,[5] and a salt bridge connecting the side chains of residues D23 and K28 has been suggested as a hallmark of Aβ aggregation.[6] Other reports have suggested that Aβ42-medi-

---

[a] *Dr. F. Attanasio,+ Dr. D. Milardi, Prof. E. Rizzarelli*
Istituto di Biostrutture e Bioimmagini-UOS CT
Consiglio Nazionale delle Ricerche
V.le A. Doria 6, 95125 Catania (Italy)
E-mail: dmilardi@unict.it

[b] *Dr. M. Convertino,+ Dr. A. Magno, Prof. A. Caflisch*
Biochemisches Institut, Universität Zürich
Winterthurerstrasse 190, 8057 Zürich (Switzerland)

[c] *Dr. A. Corazza,+ Dr. H. Haridas, Prof. G. Esposito*
Dipartimento di Scienze Mediche e Biologiche, Università di Udine
Piazzale Kolbe 4, 33100 Udine (Italy)

[d] *Dr. S. Cataldo, Prof. B. Pignataro*
Dipartimento di Fisica e Chimica, Università di Palermo
Viale delle Scienze, Parco d'Orleans II, Ed. 17, 90128 Palermo (Italy)

[e] *Prof. E. Rizzarelli*
Dipartimento di Scienze Chimiche, Università degli Studi di Catania
Viale A. Doria 6, 95125 Catania (Italy)

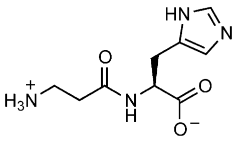
[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201200704.

ated toxicity is caused by a turn at positions E22 and D23 that favors the formation of toxic oligomers.[7]

Thus far, a number of diverse compounds have been used to prevent or reduce the aggregation of Aβ, such as curcumin,[8] Congo Red derivatives,[9] antibodies,[10] osmolytes,[11] and peptidic β-sheet breakers,[12–14] but with limited success and often with dramatic side effects. Targeting amyloid toxicity by stimulating the endogenous protective responses of the cell might represent an attractive alternative, because this would make use of mechanisms that have evolved to specifically deal with amyloid toxic assemblies.

Carnosine (β-alanyl-L-histidine) is a naturally occurring dipeptide that is present in the muscle and brain tissues of humans and other vertebrates at relatively high concentrations (1–20 mM).[15,16] Although its physiological role is not fully established, this compound has been postulated to act as a proton buffer, metal chelator, antioxidant, and as an antiglycating and antiaggregating agent.[17–22] In vitro and in vivo studies have revealed that carnosine can exert neuroprotective effects through various mechanisms. For example, it has been shown to protect neuronal cultures against glutamate-induced toxicity.[23,24] Carnosine has also been reported to protect PC12 cells from oxygen–glucose deprivation, thanks to its anti-oxidative properties.[25,26] Interestingly, carnosine has been shown to prevent Aβ-amyloid aggregation in rat brain endothelial cells[27] and to rescue cells from Aβ-induced neurotoxicity.[28–32] Furthermore, lower carnosine plasma levels have been found in AD patients than in age-matched controls.[33] Despite all these promising findings, however, no molecular interpretation is available as to how carnosine inhibits Aβ amyloid toxicity. Indeed, investigating the molecular mechanism of the inhibition effects of carnosine on Aβ aggregation would seem a highly promising challenge, as it might suggest novel lead compounds in anti-AD approaches.



carnosine

In the present work, scanning force microscopy (SFM), circular dichroism (CD), and thioflavin T (ThT) fluorescence have provided consistent evidence that carnosine is able to inhibit Aβ42 aggregation and amyloid growth in the test tube without affecting its conformational preferences. However, these experimental approaches cannot provide a molecular description of how carnosine inhibits Aβ42 aggregation. Thus, implicit solvent molecular dynamics (MD) simulations were carried out to shed light on the specificity and persistence of the interactions between carnosine and monomeric Aβ42. Furthermore, 2D TOCSY and 2D NOESY NMR spectra of the water-soluble fragment Aβ12-28 were also recorded (in the absence and presence of carnosine) in experimental conditions designed to minimize Aβ–Aβ interactions. Based on our MD and NMR results, as well as previous biophysical experiments[34–37] and atomistic simulations,[38,39] we focused our investigation on the Aβ region with the highest β-aggregation propensity, to obtain details about the Aβ aggregation–inhibition mechanism. To this aim, additional MD simulations were performed on eight replicates of each of the following capped heptapeptide segments: Aβ14–20, Aβ16–22, and Aβ18–24, in the absence (blank test) and presence of one single carnosine molecule. The simulations show van der Waals contacts between carnosine and several residues in the N-terminal half of Aβ, as well as transient electrostatic interactions with residues K16, E22, and K28, with concomitant perturbation of the propensity to form intermolecular hydrogen bonds in region 18–24. The experimental data and simulation results provide evidence that binding of carnosine, albeit mainly transient, reduces the self-assembly tendency of Aβ.

Results

Carnosine inhibits Aβ42 amyloid growth

Firstly, we studied whether carnosine could inhibit aggregation of Aβ42 by using a ThT assay (see the Experimental Section). The measurements were performed in 10 mM phosphate buffer (pH 7.4). Aβ42 samples were prepared and incubated without carnosine or in the presence of a 5-, 10-, 20-, or 100-fold molar excess of carnosine for four days at 37 °C. The inhibitory effect of carnosine on Aβ42 fibrillogenesis is shown in the upper panel of Figure 1.

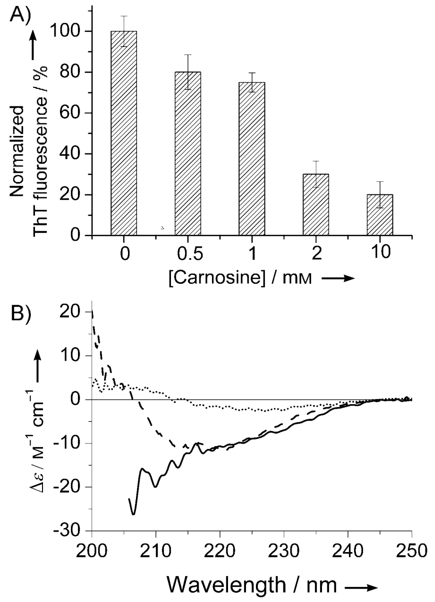


Figure 1. Top: normalized ThT fluorescence emission at 485 nm of samples containing 100 μM Aβ42 incubated at 37 °C for four days in the presence of increasing concentrations of carnosine in 10 mM phosphate buffer (pH 7.4). Bottom: CD spectra of 15 μM Aβ42 in 10 mM phosphate buffer (pH 7.4) at time $t = 0$ (——) and after incubation at 37 °C for four days in absence (⋯⋯) or presence (-----) of 0.1 mM carnosine. The dashed curve represents the difference spectrum obtained by subtracting the CD signal of carnosine from the CD curve of Aβ42 coincubated with 0.1 mM carnosine at 37 °C for four days.

The ThT results show that, at all the investigated concentrations, carnosine has a considerable effect on Aβ42 fibrillogenesis. In particular, at 20-fold molar excess carnosine exhibited a noticeable antiaggregating effect (70 % reduction in fluorescence), and an 80% reduction was observed at a 100-fold molar excess. To analyze the effect of carnosine on the conformational behaviour of Aβ42, CD experiments were carried out in the presence of the dipeptide. CD spectra of carnosine alone were also acquired under the same experimental conditions and subtracted from the Aβ42/carnosine curves. The CD spectrum of Aβ42 obtained at time $t=0$ showed that the peptide adopts a typical random-coil conformation (Figure 1 B, solid line). After four days of incubation with carnosine, the spectrum exhibited a negative minimum at 220 nm, thus suggesting the presence of a β-sheet structure (Figure 1 B, dashed line). The CD curve of pure Aβ42 (dotted line) exhibited similar spectral features but with a lower amplitude, ascribable to Aβ42 precipitation. These results confirm that the presence of carnosine does not significantly modify the conformational features of Aβ42 but, rather, interferes with its self-assembling process, thereby leading to a decrease in the total amount of fibrils formed at the end of the process. These findings were also demonstrated by SFM.

Figure 2 A shows the two types of structure that characterize the morphology of Aβ42 on mica: fibrils and globular aggregates. Magnifications of these structures are shown in Figure 2 B and C. In particular, Figure 2 B shows a group of woven fibrils about 4 nm tall and 30–40 nm wide (the latter value was corrected for the well-known tip-broadening effect; see Experimental Section). Size distribution histograms of fibril lengths and aggregate heights (Figure 2 D and E) were estimated over a large number of measurements in different regions of the sample. We observed that most of the fibrils were about 100–150 nm in length, although a small percentage reached some microns (not shown). Globular aggregates were mostly less than 1 nm in diameter (range 0.5–6.0 nm). Given that the smallest aggregate was about 0.5 nm tall (and assuming this was for one monomer, as observed by SFM), this size distribution likely includes oligomers, or packed aggregates of monomers, composed of 1 to 10 units.

Figure 3 shows the SFM analysis of the Aβ42 sample treated with carnosine. A large number of globular structures randomly dispersed on the surface are evident. Importantly, fibril aggregation was limited to a very few systems. The aggregates are much more apparent in the 3D magnification (Figure 3 B), and the related histogram (Figure 3 C) shows a monodispersed height distribution (~0.4 nm), which, as above, is consistent with the presence of monomers or small oligomers.

Interactions of carnosine with Aβ

Implicit solvent simulations were carried out on ten different systems (Table 1). At 300 K, carnosine was in contact with Aβ42 for only 20% of the simulation time (and only 10% with Aβ12–28)[40] at the concentration (5 mм) used in the simulations. Although the data at temperatures below 320 K were marred in part by incomplete convergence of sampling, it seems from analysis of the secondary structure that carnosine does not influence the overall secondary structure content of monomeric Aβ42 (Figure 4) or the secondary structure profile along the sequence (Figure S1 in the Supporting Information).

The profile of the interaction energies along the Aβ42 sequence indicates that van der Waals contacts are transient and significantly more pronounced in segment 3–22 than in the C-terminal region (Figure 5, top). Moreover, the electrostatic interactions (Figure 5, bottom) are rather promiscuous, with most of the positively and negatively charged side chains (including the terminal groups) of Aβ42 in transient contact with the carboxyl and amino group, respectively, of carnosine. Of note, the electrostatic interaction of carnosine with E22 is more pronounced than with the other negatively charged side chains (D1, E3, D7, and D23). This shows that, albeit disordered, the binding of carnosine to Aβ42 is influenced by sequence-specific effects.

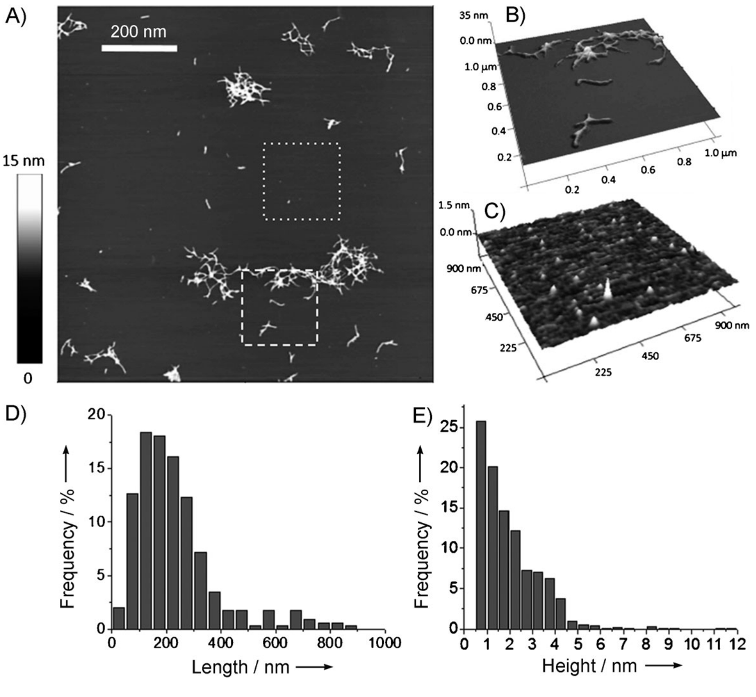


Figure 2. SFM analysis of the structures obtained by drop-casting a solution of Aβ42. A) 2D large-scale image; B) and C) 3D representations of the areas marked in (A) by dashed and dotted squares, respectively; D) length distributions of fibrils; E) height distributions of globular aggregates.

NMR characterization of Aβ12–28/carnosine contacts in solution

Ideally, to assess the presence of contacts between Aβ and carnosine by NMR the full length peptide Aβ42 should be used. However, because of its poor solubility, Aβ42 is not suitable for NMR analysis in aqueous solvent. Because of this, and the interaction profile observed in the MD simulations with Aβ42, we used the small, water-soluble fragment Aβ12–28. We also focused on this central region of Aβ42 because it encompasses residues known to be critical for Aβ aggregation.[6,41] Figure 6 depicts the overlay of the amide connectivity propagation from 2D

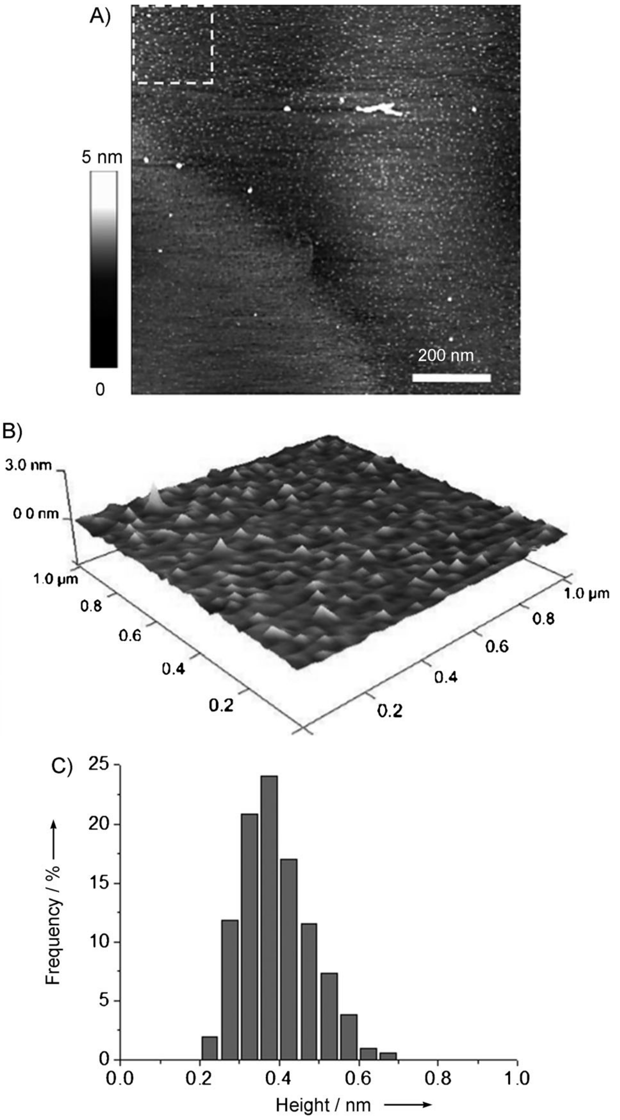


Figure 3. SFM analysis of the structures obtained by drop-casting solutions of Aβ42 and carnosine (0.1 and 10 mм, respectively). A) 2D large-scale image; B) 3D detail of the area marked in (A); C) height distribution of globular aggregates.

Table 1. Simulation parameters.

Compound	Aβ segment	Ratio	No. of runs	Length of each run [µs]	Protocol[a]	Total sampling at 300 K [µs]
carnosine	Aβ1–42	1:1	2	1.5	REMD[c]	3
–	Aβ1–42	0:1	2	1.5	REMD[c]	3
carnosine	Aβ12–28[b]	1:1	3	5	CTMD	15
–	Aβ12–28[b]	0:1	3	5	CTMD	15
carnosine	Aβ14–20	1:8	2	10	CTMD	20
–	Aβ14–20	0:8	2	10	CTMD	20
carnosine	Aβ16–22	1:8	2	10	CTMD	20
–	Aβ16–22	0:8	2	10	CTMD	20
carnosine	Aβ18–24	1:8	2	10	CTMD	20
–	Aβ18–24	0:8	2	10	CTMD	20

[a] REMD: replica exchange molecular dynamics; CTMD: constant temperature (300 K) molecular dynamics. [b] This simulation protocol is described in ref. [40]. [c] The temperatures of the individual replicas in REMD were 280, 290, 300, 310, 315, 320, 330, and 345 K.

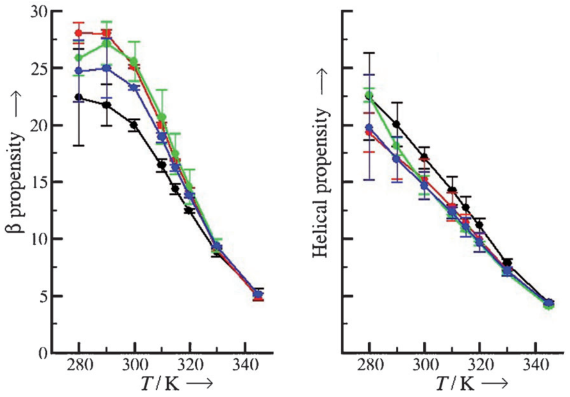


Figure 4. Secondary structure content [%] of monomeric Aβ42 with (●, ●) and without (●, ●) carnosine, as a function of REMD simulation temperature. Data are plotted separately for two independent REMD runs for each of the two systems. Error bars represent the standard deviation calculated by block averaging at each temperature and run. The β propensity does not reach convergence below 320 K, as indicated by the significant deviation between the two REMD runs in the presence of carnosine. Overall, the temperature profiles indicate a negligible influence of carnosine on the tendency of Aβ42 to form regular elements of secondary structure.

TOCSY spectra of Aβ12–28, alone and in the presence of carnosine (peptide/carnosine 1:2). For each stoichiometric equivalent addition of carnosine, the pH value increased progressively from 7.36 to 7.48, but no noteworthy shift was detected for the residues with carboxyl groups (E22, D23, and the C terminus). Chemical shift invariance with respect to carnosine additions was also observed for all the other resonances of Aβ12–28, including those of the aromatic hydrogens; this suggests lack of a stable interaction between the two molecules, or the existence of transient interactions not detectable within the timescale of NMR experiments. However, to better characterize the effect of carnosine on the peptide, the possible involvement of Aβ12–28 amide protons in H-bonding was ascertained by measuring the corresponding chemical shift thermal coefficients ($\Delta\delta/\Delta T$) over the range 6.7–12 °C. The temperature-induced migration of the individual NH signals was reconstructed from 2D TOCSY spectra. Table 2 lists the amide thermal coefficients with respect to carnosine presence: no difference is evident. The rather large $\Delta\delta/\Delta T$ values (average (7.4 ± 1.4) and (7.6 ± 1.3) ppb per degree, without and with carnosine, respectively) are consistent with a lack of stable H-bonds. The lowest $\Delta\delta/\Delta T$ values (at E22 and segment S26–N27) could be interpreted as reflecting a local trend towards H-bond formation, but somehow reduced by carnosine. The H^{α} chemical shifts, however (including those of S26 and N27) are very close to the values observed for statistically disordered peptides,[42] except for A21 and E22 where upfield shifts by 0.10 and 0.15 ppm were measured, independent of presence of carnosine. The H^{α} chemical shift deviations ($\Delta\delta H^{\alpha}$) from the tabulated values for random structures are diagnostic for secondary structure re-

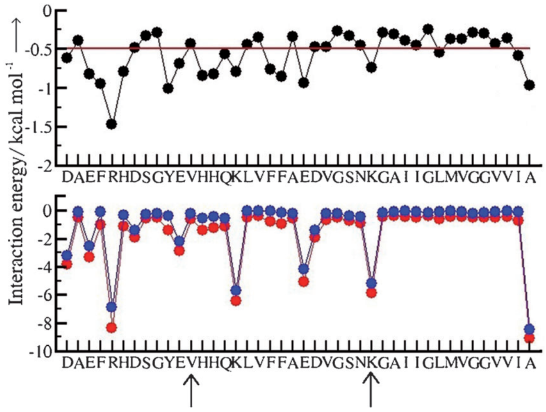


Figure 5. Interaction energy between carnosine and monomeric Aβ42 averaged over the REMD segments at 300 K (total sampling of 3 μs). The two arrows highlight the segment Aβ12–28 used in the NMR spectroscopy experiments. Top: van der Waals interaction energy. The horizontal line at −0.5 kcal mol^{-1} is a somewhat arbitrary threshold to help discriminate strongly from weakly interacting residues. Bottom: electrostatic (●) and total (●) interaction energies. The 300 K interaction energy profiles plotted separately for the two REMD runs show that these data are converged (Figure S8).

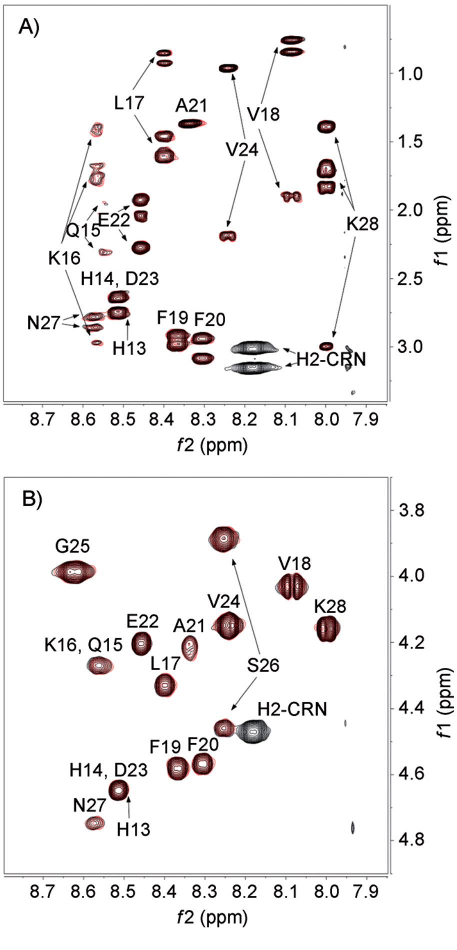


Figure 6. 2D ^{1}H TOCSY regions of 0.46 mm Aβ12–28 in H_2O/D_2O (92:8) with 4 mm NaOH and 16 mm phosphate (pH 7.36) at 6.7 °C, in the absence (red) and in the presence (black) of 0.95 mm carnosine (CRN). A) The amide connectivities to the αCHs (and βCH_2 for S26); B) the propagation to the other resonances of the spin systems. Under these conditions, excess carnosine raised the pH to 7.48, but no shift was detected for E22, D23, or K28 connectivities.

Table 2. Thermal coefficients (Δδ/ΔT) in ppb per degree of Aβ12–28[a] amide hydrogen chemical shifts in the absence and presence of carnosine[a], measured over the range 6.7–12 °C. The experimental error for all values is ±0.9 ppb per degree.

Residue	Δδ/ΔT [ppb/degree] Aβ12–28	 Aβ12–28 + carnosine
H13/H14	−6.9	−7.2
Q15/K16	−8.0	−8.9
L17	−9.1	−9.8
V18	−8.8	−8.8
F19	−9.4	−9.3
F20	−7.3	−7.4
A21	−6.6	−6.9
E22	−6.2	−6.5
D23	−6.9	−7.2
V24	−9.0	−8.8
G25	−7.0	−7.3
S26	−5.8	−6.1
N27	−4.8	−5.5
K28	−6.8	−6.8

[a] Aβ12–28 0.46 mm in 92:8 H_2O/D_2O, 16 mm phosphate and 4 mm NaOH, at pH 7.36. On addition of 0.95 mm carnosine, pH increased to 7.48. For the unresolved H13/H14/D23 and Q15/K16 amide crosspeaks (Figure 6), single Δδ/ΔT values were determined.

gions when not below 0.1 ppm and when occurring as consecutive groups.

However, based on the evidence from NOESY data (Figure 7), it is possible to attribute a structural meaning to the upfield deviations detected for A21 and E22: bending into a local α-like conformation that introduces a loose turn geometry. To investigate the possibility of peptide–carnosine association phenomena, we performed DOSY measurements in aqueous phosphate.

The results of these experiments are reported in Table 3 in terms of the hydrodynamic radii (R_h) of Aβ12–28 and carnosine, calculated according to a procedure reported elsewhere.[43] From these data, we can confirm that, with or without carnosine, Aβ12–28 is predominantly monomeric in aqueous phosphate at neutral pH and low temperature.

The apparent R_h value of carnosine increases in the presence of Aβ12–28 because of the transient fast interaction. Although less pronounced, the Aβ12–28 diffusion coefficient (and hence the apparent R_h value) is also affected by the fast intermolecular interaction. Sample aging did not alter significantly the monomeric state of Aβ12–28, at least over three to four weeks, independently of the presence of carnosine. In fact, the R_h values are constant.

The monomeric state of Aβ12–28 is not affected by temperature. In fact, no aggregation due to hydrophobic interactions (in principle favored with increasing temperature) can be inferred from the R_h values at higher temperatures. Rather, a decrease in these values was observed at higher temperatures, which should be determined by fast conformational dynamics resulting in changes in the average hydrodynamic dimensions of Aβ12–28. The failure to observe all the amide resonances in the spectra acquired at T ≥ 25 °C (because of the fast solvent exchange expected above neutral pH) is consistent with a stat-

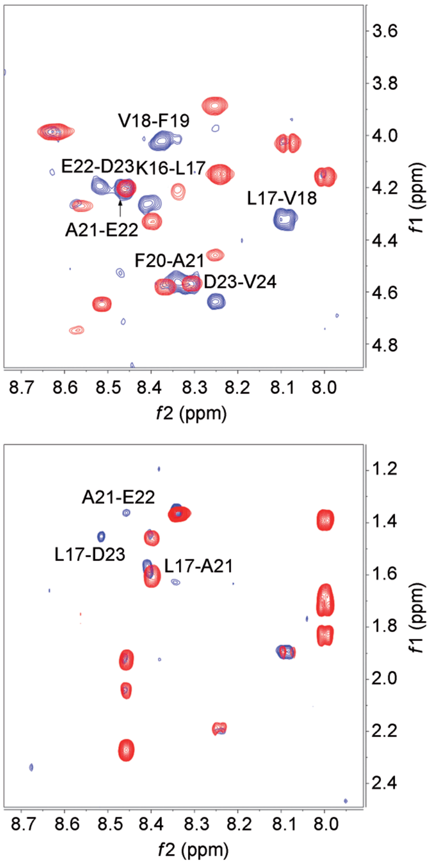


Figure 7. Overlay of 2D TOCSY (red) and 2D NOESY (blue) regions of 0.46 mM Aβ12–28 in H_2O/D_2O (92:8) with 4 mM NaOH and 16 mM phosphate (pH 7.36) at 6.7 °C. A) The intense Hα–HN connectivities; B) Hβ–HN connectivities. It is worth noting that all these connectivities occur within segment L17–V24.

Table 3. Hydrodynamic radii (R_h) of Aβ12–28[a] and carnosine measured as a function of temperature and sample age.

T [°C]	R_h Aβ12–28 [nm]	R_h carnosine [nm]	Sample age [d]
6.7	1.04 ± 0.09	–	1
6.7	1.04 ± 0.06	0.38 ± 0.02	1
6.7	–	0.34 ± 0.01	1
6.7	0.97 ± 0.05	–	7
6.7	0.98 ± 0.06	0.40 ± 0.03	7
6.7	0.99 ± 0.03	–	20
6.7	1.01 ± 0.03	0.40 ± 0.03	20
12	0.96 ± 0.01	–	21
12	0.99 ± 0.02	0.37 ± 0.01	21
25	0.92	–	28
25	0.92	0.37	28
30	0.93	–	28
30	0.93	0.37	28

[a] Aβ12–28 0.46 mM in 92:8 H_2O/D_2O, 16 mM phosphate, and 4 mM NaOH, at pH 7.36. When 0.95 mM carnosine was present, the pH increased to 7.48. The R_h values were determined from the average diffusion coefficients determined by single exponential fitting of different Aβ12–28 methyl signals and the β-alanine methylene signals of carnosine. R_h uncertainties are the standard deviations calculated from the dispersion of the diffusion coefficients. The values at 25 and 30 °C were obtained from the diffusion coefficients estimated directly from the corresponding DOSY maps. In this case uncertainty should be 5–10 %, as the R_h of TSP (the NMR chemical shift internal reference) was 0.29 nm at both temperatures, thus ensuring the reliability of the determinations.

istical conformational averaging. In accordance with the NMR analysis, MD simulations of the Aβ12–28 peptide evidenced that carnosine interacts mainly with the E22, K16, and K28 side chains of Aβ12–28, and that all these interactions are short-lived[40] (Figures S2 and S3). The decomposition of intermolecular energy into contributions of individual pairs of functional groups illustrates that the main interactions involve formal charges of opposite sign, that is, the N-terminal amino group of carnosine with E22 and D23 side chains, and the C-terminal carboxy group of carnosine with K16 and K28 side chains.

Molecular dynamics simulations of octameric Aβ heptapeptides with carnosine

The following analysis is based on MD simulations at 300 K of six systems: octameric Aβ14–20, octameric Aβ16–22, and octameric Aβ18–24, each of these three systems with and without one molecule of carnosine (Table 1).

The total sampling for each of the octameric heptapeptide system was 20 μs. A comparison of the distribution of the number of interpeptide backbone hydrogen bonds in the MD simulations with and without carnosine shows that the dipeptide reduces the aggregation tendency of the octameric Aβ18–24 system significantly, but only marginally for octameric Aβ14–20 and Aβ16–22 (Figure 8, top, see also the time series in Figures S4 and S5). The number of intrapeptide backbone hydrogen bonds increased significantly only for the Aβ18–24 heptapeptide (Figure 8 F) which is consistent with the decrease in interpeptide hydrogen bonds. The change in the distribution of the interpeptide interaction energy (Figure 8, bottom) confirms that carnosine has a minor effect on the mainly β-sheet aggregates of the octameric Aβ14–20 and Aβ16–22 systems while it has a stronger influence, that is, larger shift to less favorable values, for octameric Aβ18–24. Note that the distributions shown in Figure 8 are not affected by statistical error as shown by block averaging in Figure S6. The short residence time of carnosine on Aβ42 and on the Aβ heptapeptide systems indicate that the intermolecular interactions are weak and exhibit a transient nature (see Figures S3, S4 and S7).

Discussion

Carnosine, a naturally occurring dipeptide, is becoming a clinically accepted nutritional supplement with uses across a considerable spectrum of chronic diseases, from senile cataract to dementia. It is found in excitable tissues such as the myocardium, skeletal muscles and brain, but it is normally destroyed by the enzyme carnosinase. Normally, carnosinase acts fast and reduces the serum levels of carnosine, thus diminishing its effectiveness in clinical treatments.[21] There have been attempts at stabilizing carnosine against carnosinase degradation, for example, by conjugating it with vitamin E derivatives[44] or small sugars,[45,46] but research in this area is still underway. It is

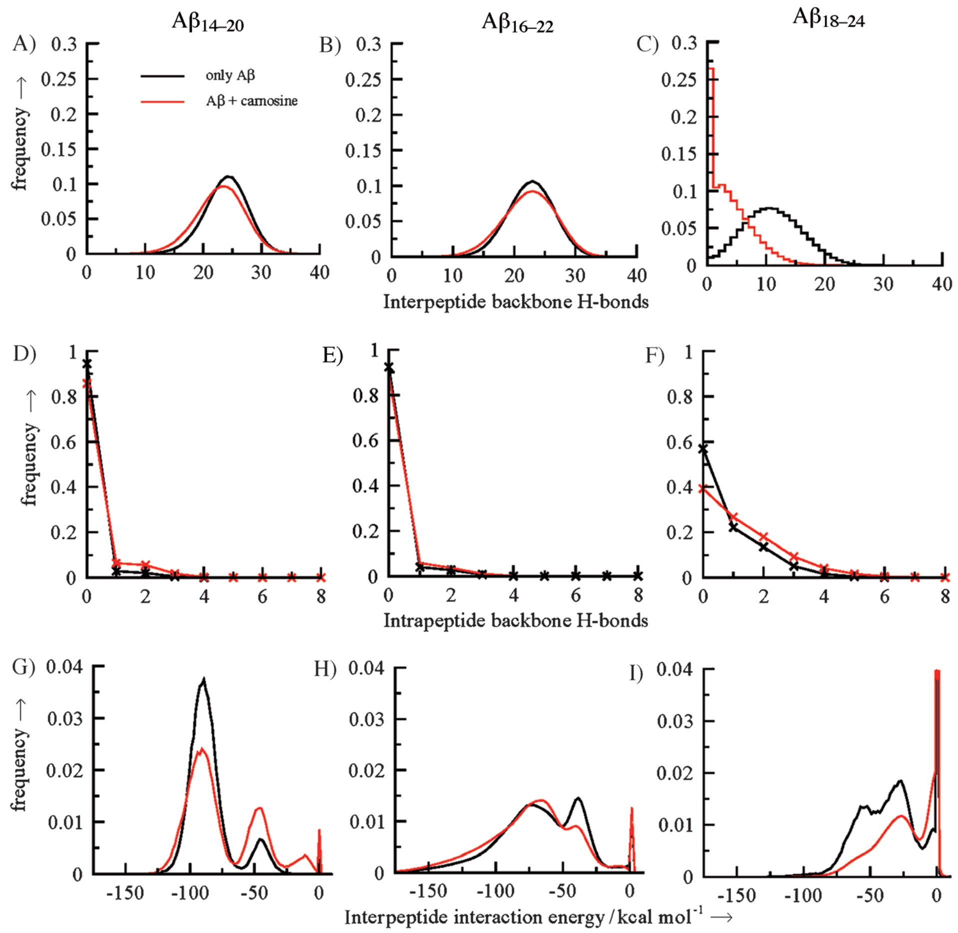


Figure 8. Influence of carnosine on the interpeptide hydrogen bonds (top graphs), intrapeptide hydrogen bonds (middle), and interpeptide interaction energy (bottom) in the octameric Aβ heptapeptide systems. Each distribution was calculated from the complete sampling (20 μs). The statistical error is much smaller than the difference between the distributions with and without carnosine (see Figure S6).

known that carnosine can protect cells exposed to high levels of Aβ42 and it has been postulated that the mechanism for protection might lie in its anti-glycating, antioxidant and/or metal-chelating activities. Inspired by the consistent body of evidence supporting a neuroprotective activity of carnosine,[47] we show here that millimolar concentrations of this dipeptide are able to effectively hinder Aβ growth in tube tests. This property of carnosine could be particularly important as direct interference with Aβ aggregation is a major target in AD therapeutics,[48] especially as a nutraceutical antagonist of Aβ self-assembly[49] has recently reached the phase II clinical trial step. On the other hand, the recent failure of the small molecule tramiprosate in phase III clinical trials calls for novel inhibitors with alternative mechanisms of modulation of Aβ aggregation.[48] Carnosine/Aβ interactions are weak and characterized by short-lived contacts localized around the central hydrophobic cluster of Aβ. These interactions do not significantly affect the conformational preferences of Aβ42. Consistent with biophysical experiments and MD simulations, NMR spectra of the fragment Aβ12–28 showed that the upfield shifts of the signals from A21 and E22, which are ascribable to local loose turn geometries, are not affected by carnosine. However, a closer inspection of NMR spectra collected at different temperatures could be interpreted as a local trend towards H-bond formation in E22 and in the segment S26–N27 being reduced by carnosine.

Consistent with NMR data, MD simulations of heptapeptide segments of Aβ show a significant decrease of the number of interpeptide H-bonds in the region 18–24. Notably, a toxic conformer of Aβ42 with a turn at positions 22 and 23 (the toxic turn) has recently been identified by solid-state NMR, and it was demonstrated that a monoclonal antibody against the toxic turn mainly detected Aβ oligomers in the brains of AD patients.[7] Moreover, the salt bridge between the side chains of D23 and K28 is also known to play a critical role in Aβ toxicity, and agents that drive these residues towards non-salt-bridge-forming conformations are thought to have therapeutic applications in AD.[50] Our data support the hypothesis that, albeit transiently, carnosine interacts with the charged groups surrounding the central hydrophobic cluster of Aβ and interferes with the network of H-bonds near the turn at positions 22 and 23, thereby hindering self-assembly. Our results, along with analogous studies, might explain in part the beneficial effect of carnosine (or derivatives thereof) in the treatment of AD.

Experimental Section

Chemicals: Aβ42 and Aβ12–28 were purchased from Anaspec (Fremont, CA). Carnosine, trifluoroacetic acid (TFA) and 1,1,1,3,3,3-hexafluoropropan-2-ol (HFIP) were from Sigma–Aldrich. The purity of all peptides was $>95\,\%$. All other chemicals were of the highest available grade and were used without further purification. Ultrapure, metal-free Milli-Q water (EMD Millipore, Billerica, MA) was used in all experiments.

Aβ peptide preparation and analysis: Monomeric Aβ42 was prepared as reported elsewhere.[51,52] The peptide was initially dissolved in TFA ($1\ \mathrm{mg\,mL^{-1}}$) and sonicated in a water bath for 10 min. Then, TFA was evaporated under a gentle stream of argon and HFIP (1 mL) was added. After incubation (37 °C, 1 h), the peptide solution was dried under a stream of argon, and then solubilized again by adding HFIP (2 mL). Finally, HFIP was removed in an argon stream, followed by drying in a lyophilizer for 1 h. The dried peptide was first dissolved in freshly prepared NaOH (2 mм), then the appropriate volume of phosphate buffer (10 mм, pH 7.4) was added to a final peptide concentration of 100 μм. Aβ42 samples without and with carnosine were incubated for four days at 37 °C.

Circular dichroism: The CD spectra of the Aβ42 sample (15 μм in phosphate buffer (10 mм, pH 7.4)) at time $t=0$ and after incubation at 37 °C for four days in the absence or presence of carnosine (0.1 mм) were recorded at 25 °C on a J-810 spectropolarimeter (Jasco), thermostated with a Jasco peltier accessory. Raw CD data acquired in a quartz cell (path length 10 mm) were processed by Spectra Manager 1.5 software (Jasco) and corrected by subtraction of the background solvent spectrum obtained under identical experimental conditions. The final CD curves represent averages of ten spectra.

ThT fluorescence measurements: Fluorescence emission spectra of ThT undergo a red shift upon incorporation into β-sheet amyloid structures. Aliquots of incubated Aβ42 peptide (100 μм) with different amounts of carnosine, were added to ThT in phosphate buffer (10 mм, pH 7.4); final concentrations: 30 μм ThT, 10 μм Aβ42. Fluorescence emission spectra were recorded in a quartz cell (light path of 1 cm) by an LS55 spectrofluorimeter ($\lambda_{ex}=440$ nm; PerkinElmer). Excitation and emission bandwidths were set to 5 nm. ThT spectra were corrected by subtraction of the background solvent spectrum obtained under identical experimental conditions. Ten consecutive spectra were recorded and averaged; the experiments were performed in triplicate. Emission at 485 nm was plotted as function of carnosine concentration.

Scanning force microscopy: Samples were prepared by drop casting Aβ42 (100 μм, 5 μL) or Aβ42/carnosine (100 μм/10 mм, 5 μL) onto a freshly cleaved mica surface and incubating for 5 min. Surfaces were then rinsed with ultrapure Milli-Q water and dried with a stream of dry nitrogen. SFM imaging was performed with a Multimode/Nanoscope IIIA (Bruker, Santa Barbara, CA) with commercially available etched-silicon Vista probes (nanoScience Instruments, Phoenix, AZ). For each image, 512×512 points were collected (scan rate 1 Hz). To get dimensional data for proteins, we used mainly height data. This is because it is well-known that the SFM lateral investigation of nanostructures is affected by the "tip-broadening effect".[53] Accordingly, tip/surface convolution leads to broadening of more than 10 nm, as typically observed by conventional SFM probes for structures (e.g., proteins) that are few nanometers in size.[54,55] The following algorithm can be used to compensate for tip broadening:

$$W_{obs} = 4\sqrt{\frac{W}{2} \times R_{tip}}$$

where W_{obs} and W are the observed and real width respectively and R_{tip} is the tip radius.[56] Particle height and fiber length data were recorded by Nanoscope software and then gathered into a single data set and statistically analyzed with Origin 8 software. Distribution histograms were obtained from hundreds of elements.

Molecular dynamics: Implicit solvent simulations were performed with ten different systems (Table 1). Six simulation systems had eight copies each of the capped heptapeptides Aβ14–20, Aβ16–22, or Aβ18–24, both in absence or presence of a single carnosine molecule (8:1 ratio). The remaining four systems consisted of monomeric Aβ42 or monomeric Aβ12–28 without or with carnosine (1:1 ratio). All Aβ peptide segments were capped by acetyl and *N*-methylamide groups at the N- and C-terminal residues, respectively, to avoid spurious charges not present in full length Aβ42. All simulations were performed with the CHARMM program;[57,58] peptides and carnosine were modeled by using the united atoms CHARMM PARAM19 force field with its default truncation scheme for nonbonding interaction (cutoff 7.5 Å). Protonation states of titratable residues were considered at pH 7.0; in particular, the imidazole of carnosine was neutral (protonated at Nδ) while its amino and carboxy groups were treated as positively and negatively charged, respectively. The electrostatic contribution to solvation was accounted for by using FACTS, an efficient generalized Born implicit solvent model, which is based on the fully analytical evaluation of the volume and spatial symmetry of the solvent that is displaced from around a solute atom by its neighboring atoms.[59] The nonpolar contributions to solvation energy was approximated by a term proportional to the surface of the solute, by using a surface-tension-like, multiplicative parameter (0.0075 kcal mol^{-1} Å^{-2}). The simulations were prepared by initially placing eight monodispersed replicates of the same heptapeptide (for the octameric systems) or one Aβ42 or Aβ12–28 peptide (for the monomeric systems), with or without a single carnosine molecule in the simulation box. Note that the initial relative positions and orientations are irrelevant because of the length of the individual trajectories (each of several microseconds), which yielded multiple association/dissociation events. Simulations were carried out with a periodic boundary condition and at a fixed peptide concentration (4.88 mм; the simulations boxes were set to 132 Å for Aβ14–20, Aβ16–22, and Aβ18–24, and to 69 Å for Aβ42 and Aβ12–28) with a Langevin integrator at low friction (coefficient 0.15 ps^{-1}). The octameric heptapeptide systems were simulated at 300 K, which yielded reversible aggregation. The same temperature was used for the Aβ12–28 segment. The simulations of full length Aβ42 with and without carnosine were conducted with the Replica Exchange Molecular Dynamics (REMD) technique at the temperatures listed in Table 1. Every 20 ps, states *i* and *j* from neighboring temperatures were swapped according to a Metropolis-like algorithm; this yielded an average acceptance ratio of 40 % for temperature swaps. For each octameric system, ten 2 μs runs were carried out, starting from randomly placed heptapeptides without any intermolecular interaction. Analogously, simulations in the presence of carnosine were started without any contact between carnosine and Aβ (full length or segment). The cumulative MD sampling at 300 K was 156 μs (198 μs when considering REMD sampling at different temperatures; Table 1).

Analysis of trajectories: The interaction energy of carnosine with Aβ42 was calculated with CHARMM, with the same threshold as for the simulations (7.5 Å). The program Wordom[60,61] was employed to calculate secondary-structure propensities according to the DSSP algorithm.[62] Carnosine influence on the ordered aggregation of octameric Aβ14–20, Aβ16–22, and Aβ18–24 was analyzed by the average number of inter- and intrapeptide H-bonds and their distributions along the simulations for each studied system. The following hydrogen bond criteria were employed: distance threshold 2.5 Å for the H–O distance, and cut off >130° for the NH–O angle. The carnosine effect on Aβ heptapeptide aggregation was estimated by calculating the interpeptide interaction energy, which is the CHARMM nonbonded energy (van der Waals plus electrostatic) of a single Aβ peptide with the remaining seven peptides, without considering the carnosine molecule.

NMR measurements: All NMR experiments were carried out at 11.7 T in an Avance 500 spectrometer (Bruker) equipped with triple-axis magnetic field gradients, by observing the hydrogen spectrum at 500.13 MHz. To avoid Aβ12–28 aggregation and precipitation, the peptide was first dissolved under basic conditions (2.3 mм in NaOH (20 mм), pH > 10).[63] Fivefold dilution in phosphate buffer led to final concentrations of peptide, phosphate, and NaOH of 0.46, 16, and 4 mM, respectively, in H_2O/D_2O (92:8, pH 7.36). Carnosine additions were performed with aliquots (5–10 μL) from a stock solution in phosphate buffer (50 mM, pH 7.42)) to reach stoichiometric ratios 1:1 and 2:1 (carnosine:Aβ12–28), with a final pH of 7.48. To inhibit hydrophobic interaction, all operations were performed in an ice bath, and solutions were stored at 4 °C. For NMR measurements, the temperature was set initially to 6.7 °C and subsequently increased stepwise: 6.7–12 °C, then 25–30 °C. Chemical shifts were referenced to internal 3-trimethyl-silyl 2,2,3,3-tetradeutero sodium propionate (TSP). In addition to 1D spectra, 2D TOCSY[64] and NOESY[65] experiments were performed, typically with 2048×512 (t2×t1) data-point matrices for spectral widths of 6 kHz in each dimensions with 64 scans per t_1 point, 64 dummy scans, and relaxation delay of 1 s. The TOCSY MLEV17[66] isotropic mixing intervals were 80 ms at $\gamma B_2/2\pi = 8.3$ kHz; for NOESY, mixing times of 250 ms were employed. Quadrature in the indirect dimension was obtained by time-proportional phase incrementation,[67] and solvent suppression was implemented by appending a WATERGATE spin-echo module to the sequences, performed in the excitation-sculpting mode.[68,69] This solvent suppression scheme was also used in diffusion ordered spectroscopy (DOSY)[70] experiments for the measurement of the diffusion coefficients. A sequence including bipolar gradients and longitudinal eddy current suppression was employed in a double-stimulated echo module delay for cancellation of convection artifacts,[71] to collect 80 points with linear increments from 2 to 95 % of the z-axis magnetic field gradient (65 G cm^{-1}), with delays $\Delta = 90$ ms, $\sigma = 2\times2$ ms, and 128 scans/increment. All data were processed with TOPSPIN software (Bruker).

Acknowledgements

This work was supported by MIUR (grants: PRIN 2010M2JARJ_001; PRIN 20083ERXWS_005; FIRB-MERIT no. RBNE08HWLZ_010). M.C. and A. Caflisch are grateful to Riccardo Pellarin, Andreas Vitalis and the late François Marchand for interesting discussions. The simulations were carried out on the Schroedinger cluster of the University of Zürich. The work in Zürich was supported financially by the Swiss National Science Foundation and the Swiss National Competence Center (NCCR) on Neural Plasticity and Repair.

Keywords: Alzheimer's disease · carnosine · molecular dynamics · neuroprotective agent · nutraceutical compounds · protein–protein interactions

[1] J. Hardy, *Neurobiol. Aging* **2002**, *23*, 1073 – 1074.
[2] J. Kang, H.-G. Lemaire, A. Unterbeck, J. M. Salbaum, C. L. Masters, K.-H. Grzeschik, G. Multhaup, K. Beyreuther, B. Müller-Hill, *Nature* **1987**, *325*, 733 – 736.
[3] C. Haass, D. J. Selkoe, *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 101 – 112.
[4] M. L. Giuffrida, F. Caraci, B. Pignataro, S. Cataldo, P. De Bona, V. Bruno, G. Molinaro, G. Pappalardo, A. Messina, A. Palmigiano, D. Garozzo, F. Nicoletti, E. Rizzarelli, A. Copani, *J. Neurosci.* **2009**, *29*, 10582 – 10587.
[5] M. Ahmed, J. Davis, D. Aucoin, T. Sato, S. Ahuja, S. Aimoto, J. I. Elliott, W. E. Van Nostrand, S. O. Smith, *Nat. Struct. Mol. Biol.* **2010**, *17*, 561 – 567.
[6] K. L. Sciarretta, D. J. Gordon, A. T. Petkova, R. Tycko, S. C. Meredith, *Biochemistry* **2005**, *44*, 6003 – 6014.
[7] N. Izuo, T. Kume, M. Sato, K. Murakami, K. Irie, Y. Izumi, A. Akaike, *ACS Chem. Neurosci.* **2012**, *3*, 674 – 681.
[8] A. N. Begum, M. R. Jones, G. P. Lim, T. Morihara, P. Kim, D. D. Heath, C. L. Rock, M. A. Pruitt, F. Yang, B. Hudspeth, S. Hu, K. F. Faull, B. Teter, G. M. Cole, S. A. Frautschy, *J. Pharmacol. Exp. Ther.* **2008**, *326*, 196 – 208.
[9] P. Frid, S. V. Anisimov, N. Popovic, *Brain Res. Rev.* **2007**, *53*, 135 – 160.
[10] B. Solomon, *Curr. Opin. Invest. Drugs* **2007**, *8*, 519 – 524.
[11] D.-S. Yang, C. M. Yip, T. H. J. Huang, A. Chakrabartty, P. E. Fraser, *J. Biol. Chem.* **1999**, *274*, 32970 – 32974.
[12] P. Soto, M. A. Griffin, J.-E. Shea, *Biophys. J.* **2007**, *93*, 3015 – 3025.
[13] C. Soto, E. M. Sigurdsson, L. Morelli, R. A. Kumar, E. M. Castaño, B. Frangione, *Nat. Med.* **1998**, *4*, 822 – 826.
[14] P. De Bona, M. L. Giuffrida, F. Caraci, A. Copani, B. Pignataro, F. Attanasio, S. Cataldo, G. Pappalardo, E. Rizzarelli, *J. Pept. Sci.* **2009**, *15*, 220 – 228.
[15] R. Kohen, Y. Yamamoto, K. C. Cundy, B. N. Ames, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 3175 – 3179.
[16] J. J. O'Dowd, D. J. Robins, D. J. Miller, *Biochim. Biophys. Acta Gen. Subj.* **1988**, *967*, 241 – 249.
[17] H. Abe, *Biochemistry (Moscow)* **2000**, *65*, 757 – 765.
[18] A. R. Hipkiss, J. E. Preston, D. T. M. Himsworth, V. C. Worthington, M. Keown, J. Michaelis, J. Lawrence, A. Mateen, L. Allende, P. A. M. Eagles, N. J. Abbott, *Ann. N. Y. Acad. Sci.* **1998**, *854*, 37 – 53.
[19] F. Attanasio, S. Cataldo, S. Fisichella, S. Nicoletti, V. G. Nicoletti, B. Pignataro, A. Savarino, E. Rizzarelli, *Biochemistry* **2009**, *48*, 6522 – 6531.
[20] V. Calabrese, E. Guagliano, M. Sapienza, A. Ravagna, V. Cardile, G. Scapagnini, A. M. Santoro, A. Mangiameli, D. A. Butterfield, A. M. Giuffrida Stella, E. Rizzarelli, *Neurochem. Res.* **2005**, *30*, 797 – 807.
[21] F. Bellia, G. Vecchio, E. Rizzarelli, *Amino Acids* **2012**, *43*, 153 – 163.
[22] V. Calabrese, C. Cornelius, S. Cuzzocrea, I. Iavicoli, E. Rizzarelli, E. J. Calabrese, *Mol. Aspects Med.* **2011**, *32*, 279 – 304.
[23] A. Boldyrev, E. Bulygina, T. Leinsoo, I. Petrushanko, S. Tsubone, H. Abe, *Comp. Biochem. Physiol. Part B* **2004**, *137*, 81 – 88.
[24] A. Boldyrev, R. Song, D. Lawrence, D. O. Carpenter, *Neuroscience* **1999**, *94*, 571 – 577.
[25] R. Tabakman, H. Jiang, R. A. Levine, R. Kohen, P. Lazarovici, *Neurosci. Res.* **2004**, *75*, 499 – 507.
[26] F. Bellia, G. Vecchio, S. Cuzzocrea, V. Calabrese, E. Rizzarelli, *Mol. Aspects Med.* **2011**, *32*, 258 – 266.
[27] J. E. Preston, A. R. Hipkiss, D. T. Himsworth, I. A. Romero, J. N. Abbott, *Neurosci. Lett.* **1998**, *242*, 105 – 108.
[28] Q. Fu, H. Dai, W. Hu, Y. Fan, Y. Shen, W. Zhang, Z. Chen, *Cell. Mol. Neurobiol.* **2008**, *28*, 307 – 316.
[29] H. Atamna, R. Kumar, *J. Alzheimer's Dis.* **2010**, *20*, S439 – S452.
[30] X. Fernàndez-Busquets, J. Ponce, R. Bravo, M. Arimon, T. Martiáñez, A. Gella, J. Cladera, N. Durany, *Curr. Alzheimer Res.* **2010**, *7*, 428 – 438.
[31] A. R. Hipkiss, *J. Alzheimer's Dis.* **2007**, *11*, 229 – 240.
[32] A. Boldyrev, A. Koudinov, T. Berezov, D. O. Carpenter, *J. Alzheimer's Dis.* **2004**, *6*, 633 – 638.
[33] A. N. Fonteh, R. J. Harrington, A. Tsai, P. Liao, M. G. Harrington, *Amino Acids* **2007**, *32*, 213 – 224.
[34] L. O. Tjernberg, J. Näslund, F. Lindqvist, J. Johansson, A. Karlström, J. Thyberg, L. Terenius, C. Nordstedt, *J. Biol. Chem.* **1996**, *271*, 8545 – 8548.
[35] A. D. Williams, E. Portelius, I. Kheterpal, J.-t. Guo, K. D. Cook, Y. Xu, R. Wetzel, *J. Mol. Biol.* **2004**, *335*, 833 – 842.
[36] F. T. Senguen, N. R. Lee, X. Gu, D. M. Ryan, T. M. Doran, E. A. Anderson, B. L. Nilsson, *Mol. BioSyst.* **2011**, *7*, 486 – 496.
[37] F. T. Senguen, T. M. Doran, E. A. Anderson, B. L. Nilsson, *Mol. BioSyst.* **2011**, *7*, 497 – 510.
[38] M. Cecchini, R. Curcio, M. Pappalardo, R. Melki, A. Caflisch, *J. Mol. Biol.* **2006**, *357*, 1306 – 1321.
[39] A. Caflisch, *Curr. Opin. Chem. Biol.* **2006**, *10*, 437 – 444.
[40] M. Convertino, A. Vitalis, A. Caflisch, *J. Biol. Chem.* **2011**, *286*, 41578 – 41588.
[41] T. C. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohmann, H. Döbeli, D. Schubert, R. Riek, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17342 – 17347.
[42] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, *5*, 67 – 81.
[43] K. R. Harris, L. A. Woolf, *J. Chem. Eng. Data* **2004**, *49*, 1064 – 1069.
[44] S. L. Stvolinsky, E. R. Bulygina, T. N. Fedorova, K. Meguro, T. Sato, O. V. Tyulina, H. Abe, A. Boldyrev, *Cell. Mol. Neurobiol.* **2010**, *30*, 395 – 404.
[45] V. Lanza, F. Bellia, R. D'Agata, G. Grasso, E. Rizzarelli, G. Vecchio, *J. Inorg. Biochem.* **2011**, *105*, 181 – 188.

[46] G. I. Grasso, G. Arena, F. Bellia, G. Maccarrone, M. Parrinello, A. Pietropaolo, G. Vecchio, E. Rizzarelli, *Chem. Eur. J.* **2011**, *17*, 9448 – 9455.
[47] C. Corona, V. Frazzini, E. Silvestri, R. Lattanzio, R. La Sorda, M. Piantelli, L. M. T. Canzoniero, D. Ciavardelli, E. Rizzarelli, S. L. Sensi, *PLoS One* **2011**, *6*, e17971.
[48] I. W. Hamley, *Chem. Rev.* **2012**, *112*, 5147 – 5192.
[49] C. Rivière, T. Richard, L. Quentin, S. Krisa, J. M. Mérillon, J. P. Monti, *Bioorg. Med. Chem.* **2007**, *15*, 1160 – 1167.
[50] V. S. Mithu, B. Sarkar, D. Bhowmik, M. Chandrakesan, S. Maiti, P. K. Madhu, *Biophys. J.* **2011**, *101*, 2825 – 2832.
[51] B. O'Nuallain, A. K. Thakur, A. D. Williams, A. M. Bhattacharyya, S. Chen, G. Thiagarajan, R. Wetzel, *Methods Enzymol.* **2006**, *413*, 34 – 74.
[52] M. G. Zagorski, J. Yang, H. Shao, K. Ma, H. Zeng, A. Hong, *Methods Enzymol.* **1999**, *309*, 189 – 204.
[53] V. J. Garcia, L. Martinez, J. M. Briceño-Valero, C. H. Schilling, *Probe Microsc.* **1997**, *1*, 107 – 116.
[54] B. Pignataro, L. Sardone, G. Marletta, *Nanotechnology* **2003**, *14*, 245 – 249.
[55] B. Pignataro, *J. Mater. Chem.* **2009**, *19*, 3338 – 3350.
[56] J. Vesenka, M. Guthold, C. L. Tang, D. Keller, E. Delaine, C. Bustamante, *Ultramicroscopy* **1992**, *42 – 44*, 1243 – 1249.
[57] B. R. Brooks, R. E. Bruccolieri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187 – 217.
[58] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera et al., *J. Comput. Chem.* **2009**, *30*, 1545 – 1614.
[59] U. Haberthür, A. Caflisch, *J. Comput. Chem.* **2008**, *29*, 701 – 715.
[60] M. Seeber, A. Felline, F. Raimondi, S. Muff, R. Friedman, F. Rao, A. Caflisch, F. Fanelli, *J. Comput. Chem.* **2011**, *32*, 1183 – 1194.
[61] M. Seeber, M. Cecchini, F. Rao, G. Settanni, A. Caflisch, *Bioinformatics* **2007**, *23*, 2625 – 2627.
[62] W. Kabsch, C. Sander, *Biopolymers* **1983**, *22*, 2577 – 2637.
[63] Y. Fezoui, D. M. Hartley, J. D. Harper, R. Khurana, D. M. Walsh, M. M. Condron, D. J. Selkoe, P. T. Lansbury, A. L. Fink, D. B. Teplow, *Amyloid* **2000**, *7*, 166 – 178.
[64] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 521 – 528.
[65] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546 – 4553.
[66] A. Bax, D. G. J. Davis, *J. Magn. Reson.* **1985**, *65*, 355 – 360.
[67] D. Marion, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967 – 974.
[68] M. Piotto, V. Saudek, V. J. Sklenár, *J. Biomol. NMR* **1992**, *2*, 661 – 665.
[69] T. L. Hwang, A. J. Shaka, *J. Magn. Reson. A* **1995**, *112*, 275 – 279.
[70] K. F. Morris, C. S. Johnson, Jr., *J. Am. Chem. Soc.* **1992**, *114*, 3139 – 3141.
[71] A. Jerschow, N. Müller, *J. Magn. Reson.* **1997**, *125*, 372 – 375.

Received: November 13, 2012

CHEMBIOCHEM

Supporting Information

Carnosine Inhibits Aβ_{42} Aggregation by Perturbing the H-Bond Network in and around the Central Hydrophobic Cluster

Francesco Attanasio,[a] Marino Convertino,[b] Andrea Magno,[b] Amedeo Caflisch,[b]
Alessandra Corazza,[c] Haritha Haridas,[c] Gennaro Esposito,[c] Sebastiano Cataldo,[d]
Bruno Pignataro,[d] Danilo Milardi,*[a] and Enrico Rizzarelli[a, e]

cbic_201200704_sm_miscellaneous_information.pdf

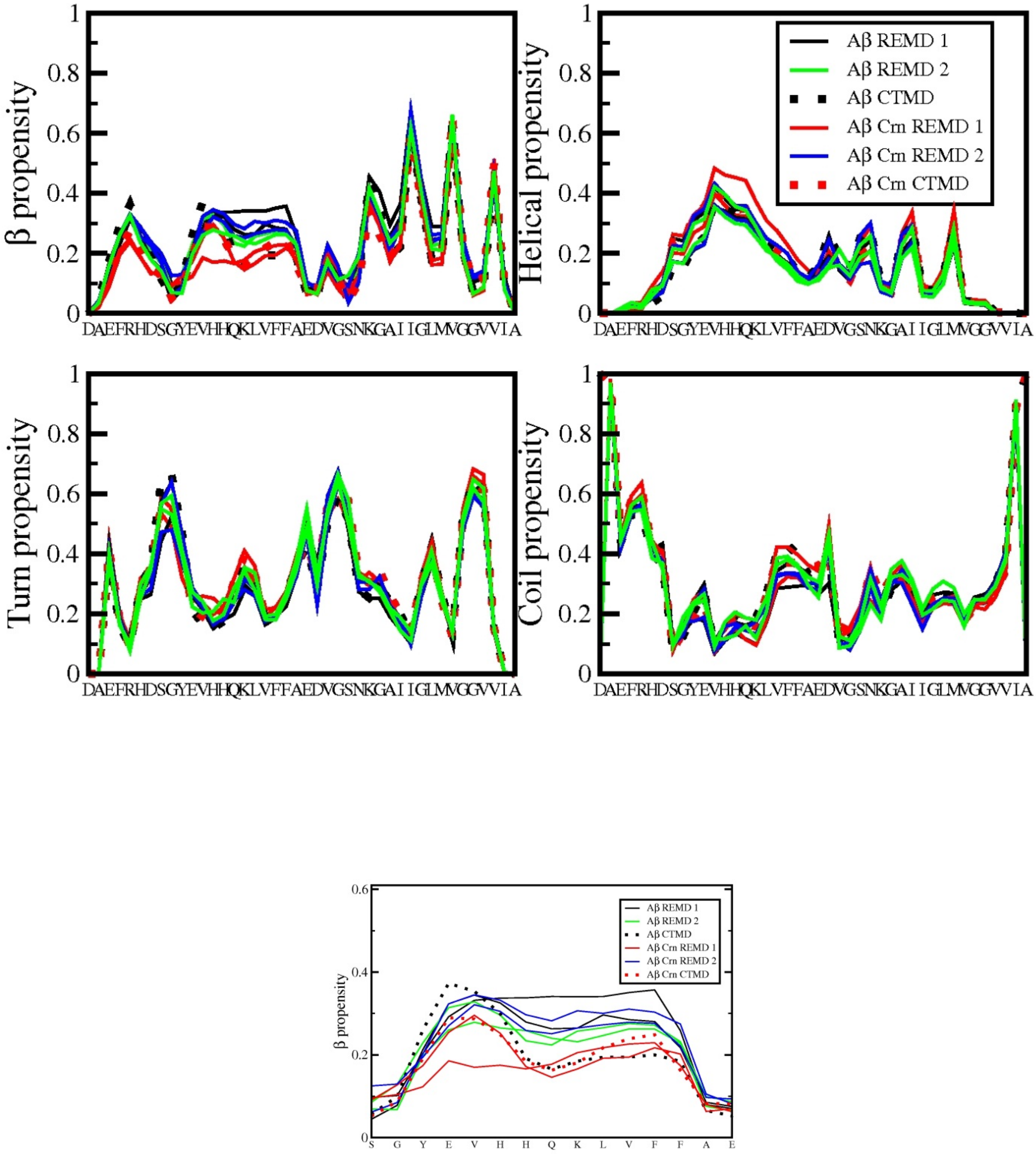
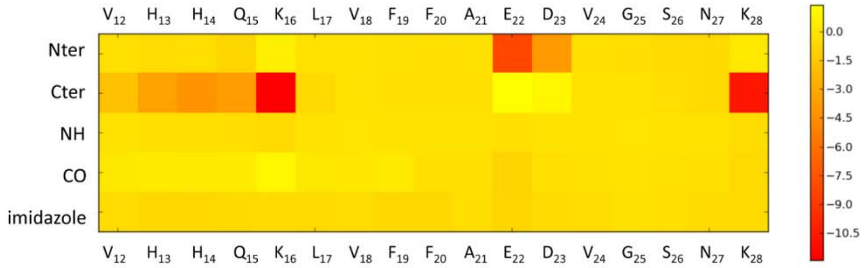
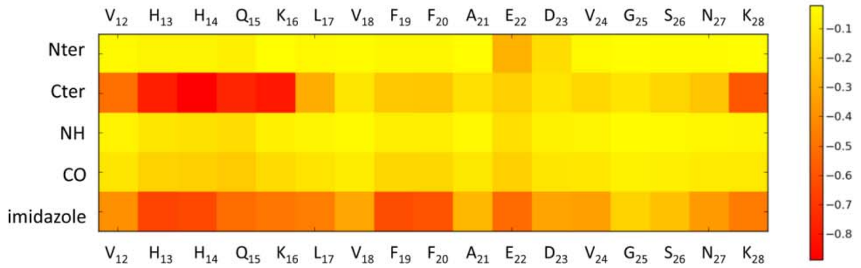


Figure S1. Secondary structure propensities along the Aβ42 sequence at T = 300K, with and without carnosine. For each system, two independent simulations have been performed with the REMD technique (solid lines). Moreover, a constant temperature molecular dynamics (CTMD, dashed lines) simulation has been run at 300 K to increase confidence in the statistical robustness of the results. The panel in the bottom shows an enlargement of the profile with the highest statistical noise which is the β-propensity of the segment from Ser8 to Glu22.

total interaction energy (kcal/mol)



van der Waals interaction energy (kcal/mol)



electrostatic interaction energy (kcal/mol)

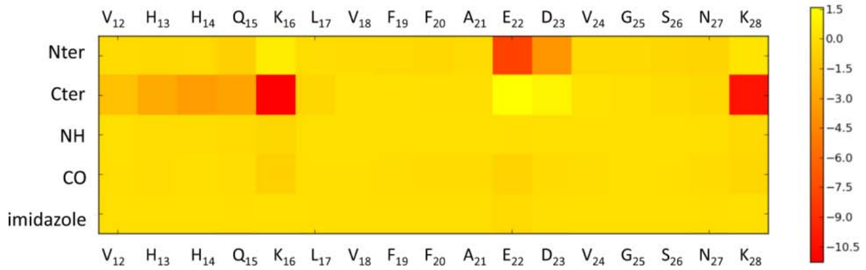


Figure S2. Interaction energy matrix for carnosine and monomeric Aβ12-28. Each colored square shows the interaction energy between a single functional group of carnosine and a single Aβ12-28 residue (backbone and side chain atoms) averaged over the whole 15 microsecond sampling. The decomposition of intermolecular energy into contributions of individual pairs of functional groups illustrates that the main interactions involve formal charges of opposite sign, i.e., the N-terminal amino group of carnosine with the E22 and D23 side chains, and the C-terminal carboxy group of carnosine with the K16 and K28 side chains. Although the electrostatic contribution seems to dominate, one has to underline that desolvation effects are not taken into account in the interaction energy matrix. The electrostatic desolvation of charged groups is unfavorable and does significantly counter-balance the favorable electrostatic interactions between the termini of carnosine and the charged side chains of Aβ. There are also favorable van der Waals interactions involving the imidazole ring of carnosine with mainly the aromatic side chains of Aβ but also other Aβ residues, as well as the carboxy group of carnosine and the Aβ12-16 segment.

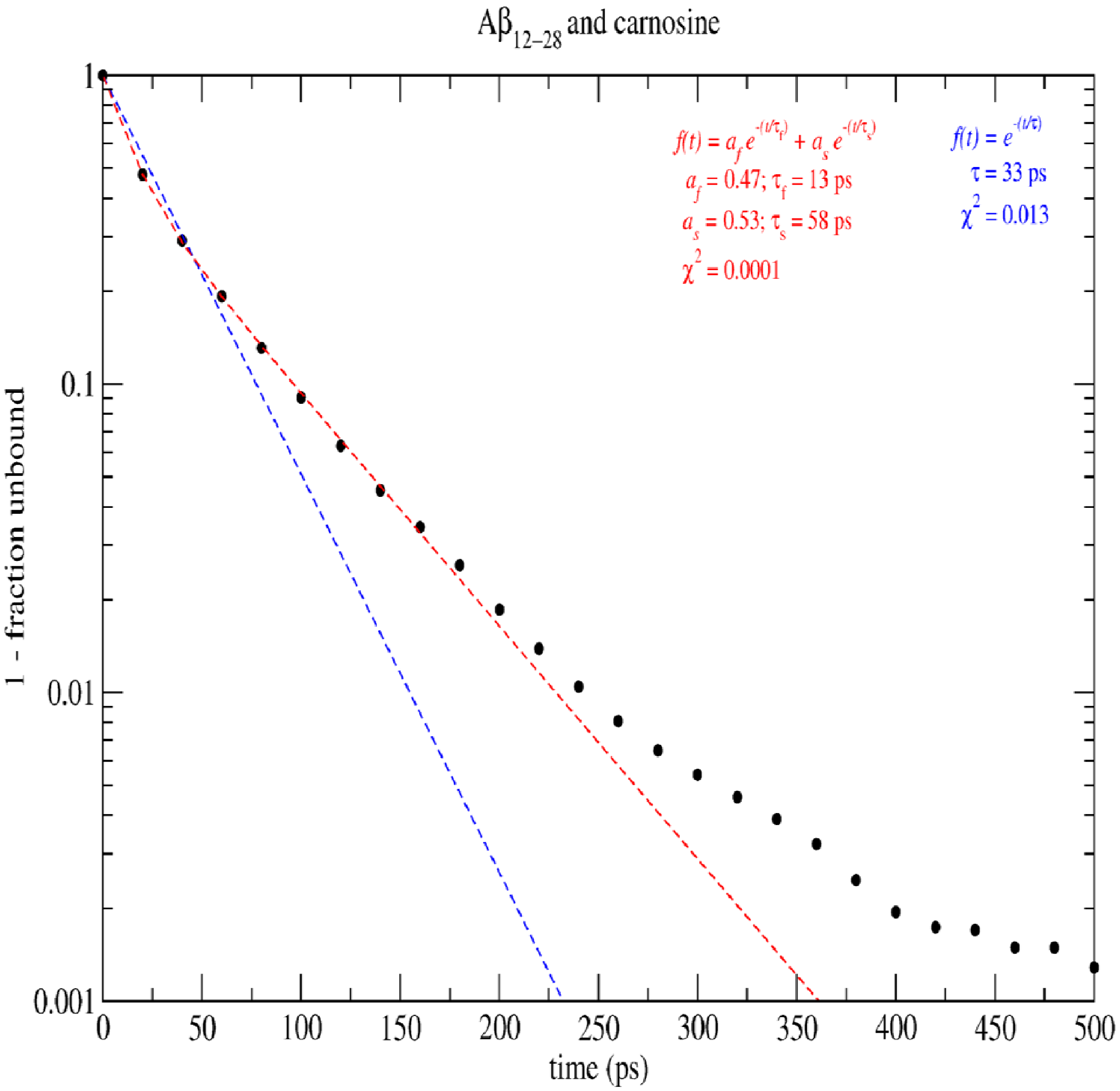


Figure S3. Cumulative distribution of the time of carnosine unbinding from Aβ12-28 extracted from the MD trajectories (filled circles). The kinetics of unbinding show a negligible dependence of the energy threshold used to define a binding event in the range from -20 to 0 kcal/mol. The cumulative distribution of the time of unbinding of carnosine from Aβ12-28 can be fitted by a double-exponential (red dashed line and legend) with similar amplitude for the slow and fast phase. This result and the characteristic time of only 58 ps for the slow phase indicate that the intermolecular interactions are transient (fast unbinding). The transient nature of the association of carnosine to monomeric Aβ12-28 is consistent with the increase of the hydrodynamic radius (R_h) of carnosine and the minor increase of the R_h of Aβ12-28 as derived from the NMR DOSY measurements. Note that the unbinding times of carnosine represent lower bounds because of the low friction used in the Langevin simulations, and it is likely that the real unbinding times are 10-100 times slower as suggested by previously reported comparisons of implicit solvent simulations with experimental data on peptide folding times (Hiltpold et al. 2000). In any case, even upon correcting for the lack of friction, unbinding time scales of about 0.5 ns to 5 ns still correspond to weak/short lived association.

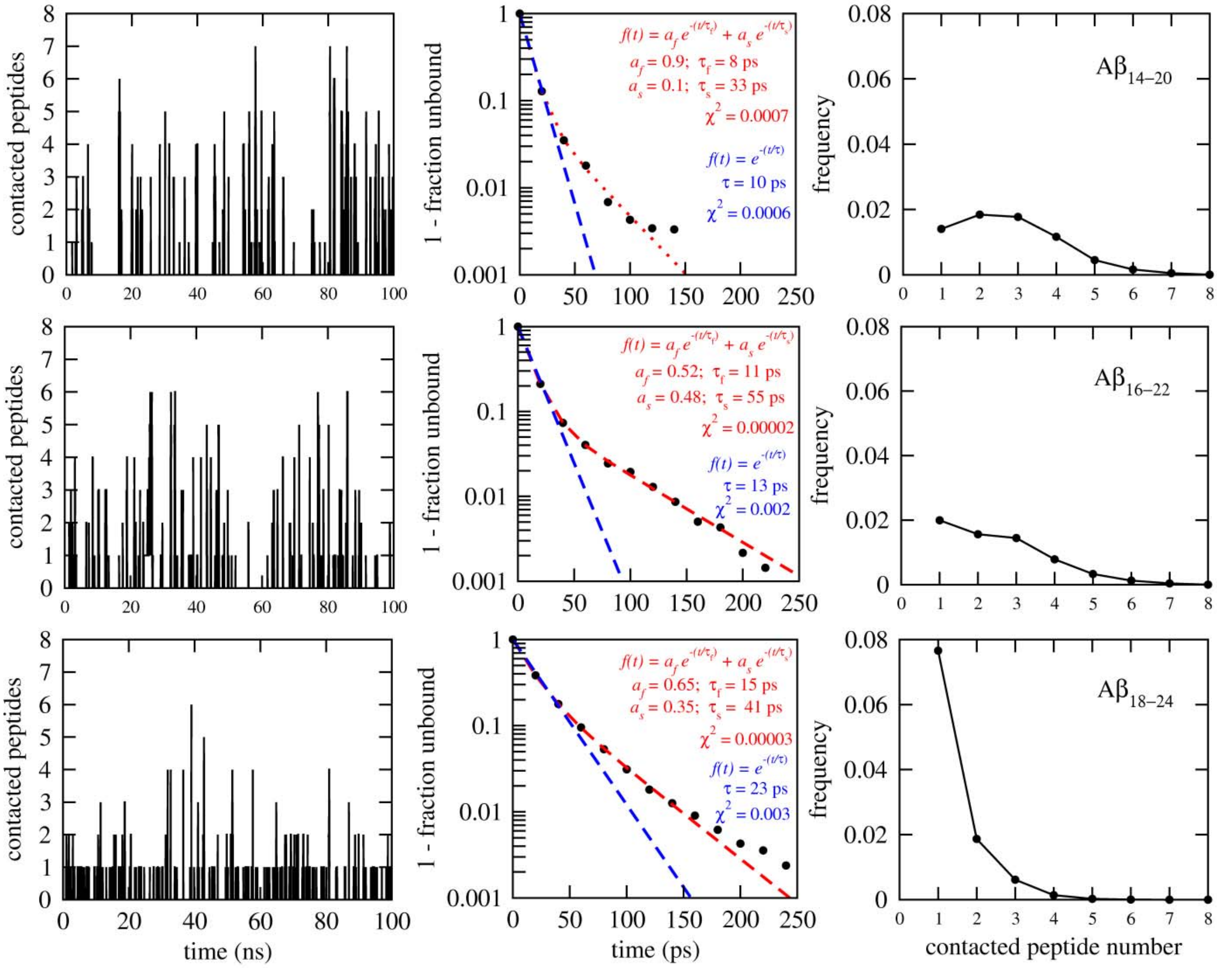


Figure S4. Transient binding of carnosine to oligomers of Aβ14-20 (top), Aβ16-22 (middle), and Aβ18-24 (bottom). The figure shows the time series (left column) and distribution (right column) of the number of Aβ heptapeptides in contact with carnosine. The time series show a trajectory subsegment of only 100 ns for clarity. The distributions are normalized to one by including in the count fully isolated carnosine whose data point (at x=0) is not shown as it lies outside the y-axis range. (Middle column) The cumulative distribution of the time of carnosine unbinding from the Aβ heptapeptide oligomers is shown with fitting curves and parameters in red and blue for double-exponential and single-exponential fitting, respectively. The residence time of carnosine on the oligomeric Aβ heptapeptide systems can be visualized by the time series of the number of Aβ heptapeptides in contact with carnosine (left). The double exponential fitting yields characteristic times of about 10 ps and 50 ps. Visual analysis of the trajectories of Aβ14-20 and Aβ16-22 indicates that the slow phase originates from carnosine dissociation from aggregates with significant intermolecular β-sheet structure, with which carnosine is involved in multiple transient interactions.

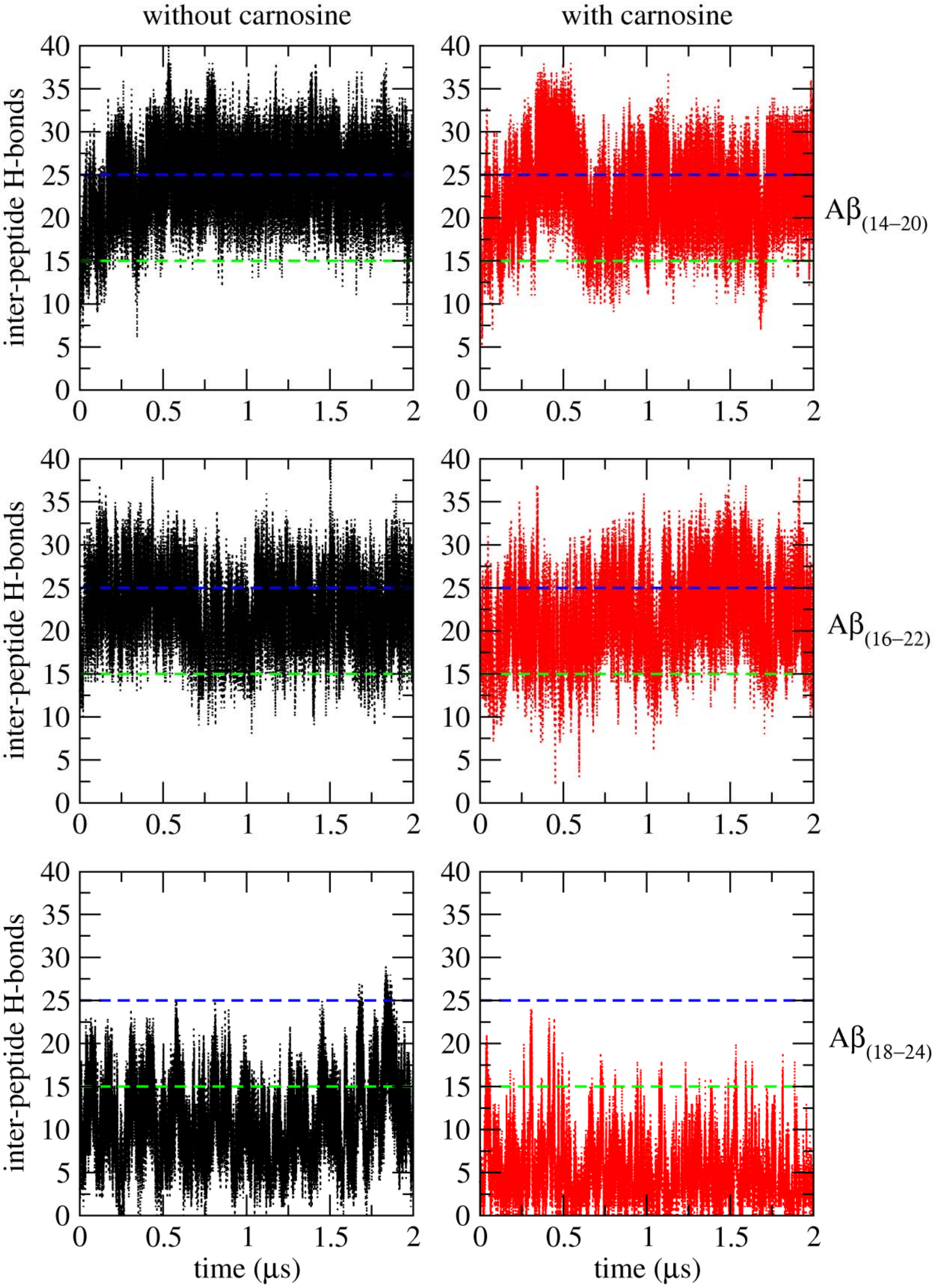


Figure S5. Time series of number of inter-peptide backbone hydrogen bonds. The horizontal lines at y=15 and y=25 are drawn to facilitate the comparison of the different plots.

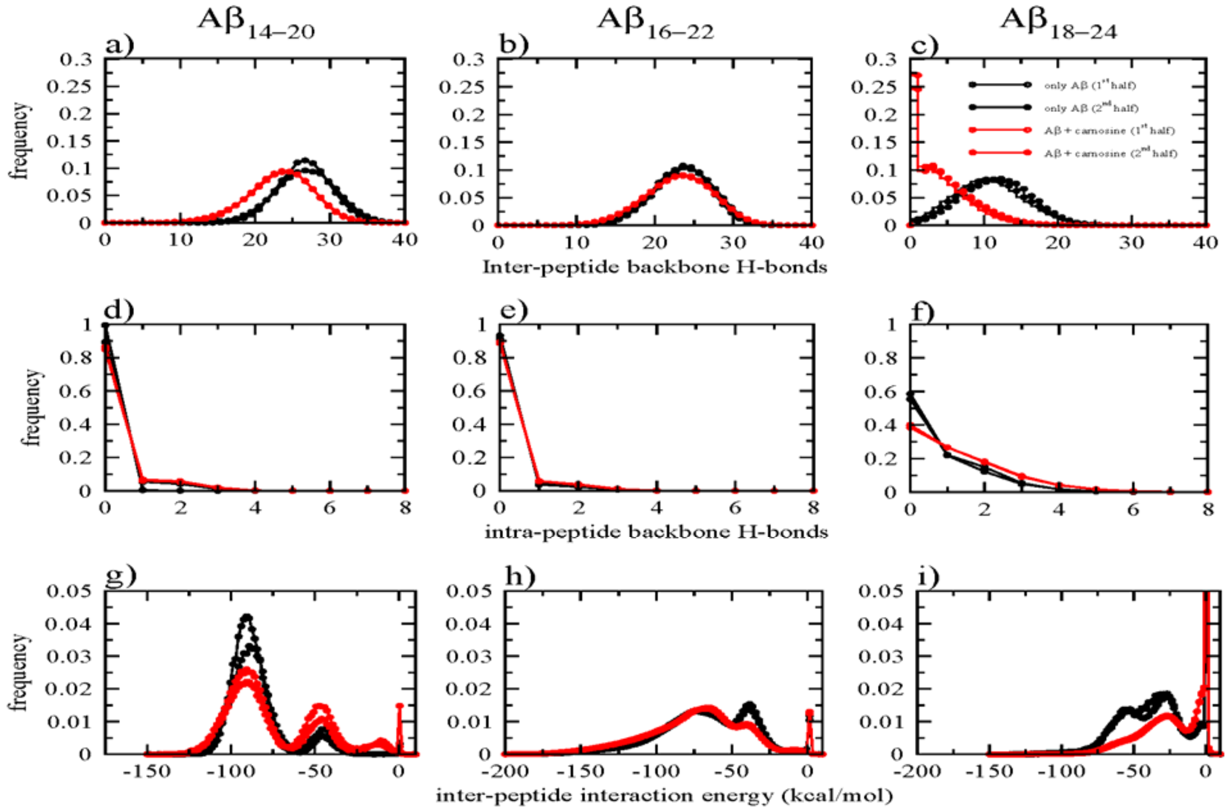


Figure S6. Distributions of number of intermolecular hydrogen bonds (top), intramolecular hydrogen bonds (middle), and intermolecular energy (bottom). For each system, block averaging was carried out by dividing the total sampling of 20 µs (ten 2-µs runs) into two blocks of 10 µs each (five MD runs). The differences in the presence and absence of carnosine (compare red with black lines) are significantly larger than the statistical noise (compare lines of same color). Moreover, the almost indistinguishable histograms of the two blocks of each system indicate that the statistical error is negligible.

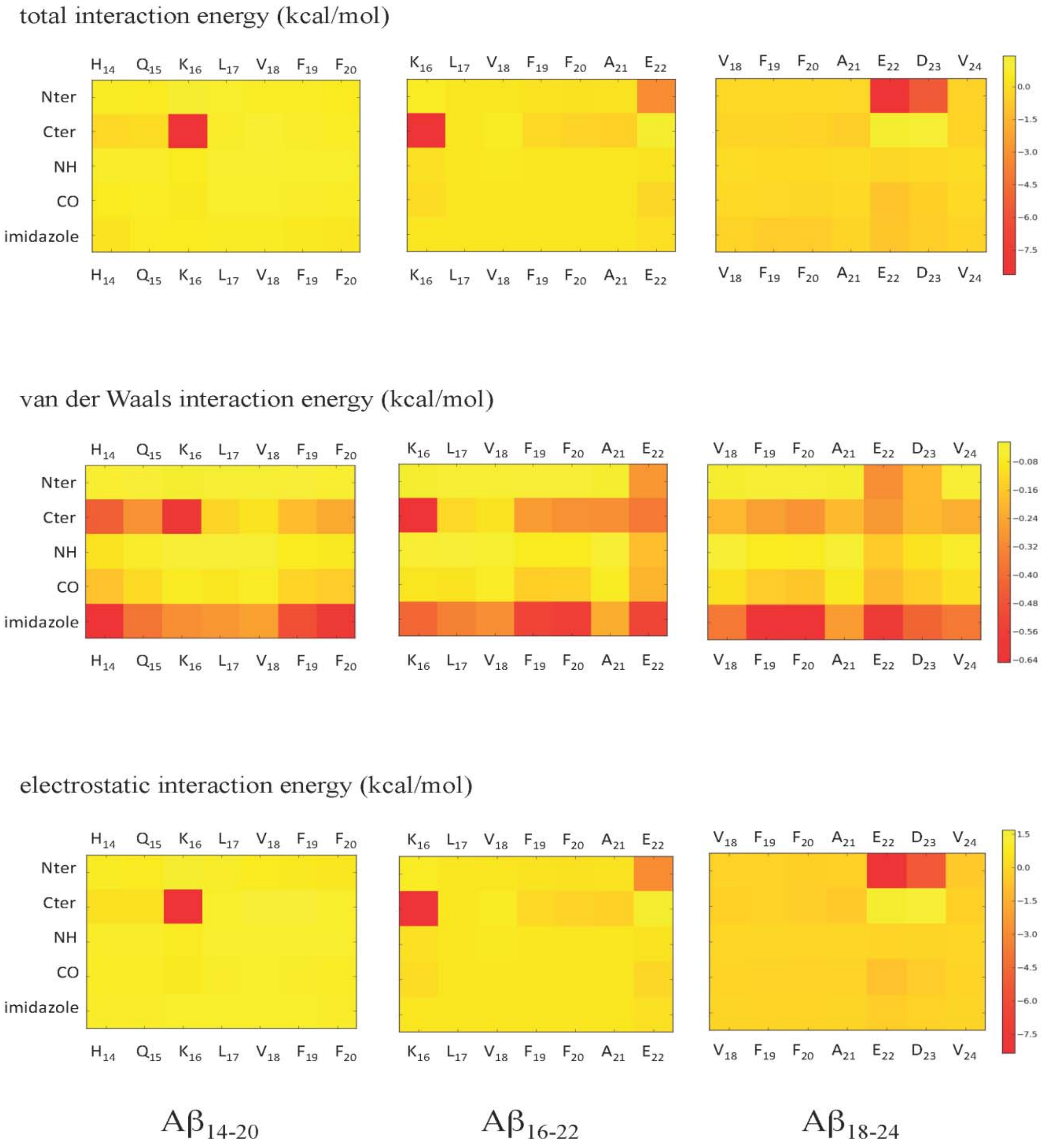


Figure S7. Interaction energy matrix for carnosine and Aβ14-20 (left), Aβ16-22 (middle), and Aβ18-24 (right).

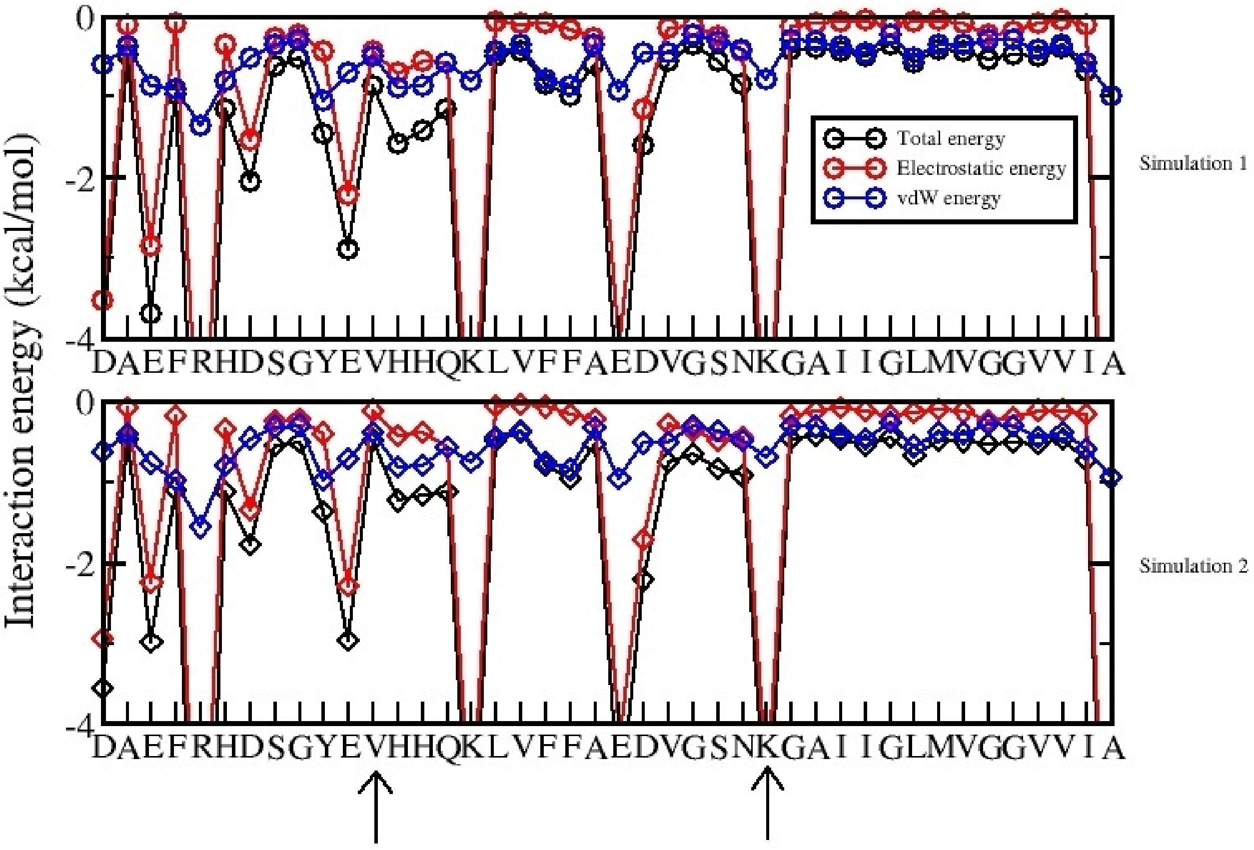


Figure S8. The 300 K interaction energy profiles of carnosine with Aβ42 plotted separately for the two REMD runs show that these data are converged. The two vertical arrows emphasize the segment Aβ12-28 used in the NMR experiments.

References

Hiltpold et al. *J. Phys. Chem. B* 2000, **104**, 10080-10086.

Chapter 6

Mechanism and kinetics of acetyl-lysine binding to bromodomains

Magno, A.[‡], Steiner, S.[‡] and Caflisch, A. *J. Chem. Theory Comput.*, **2013**, *9, 4225-4232*

[‡]These authors contributed equally to this work

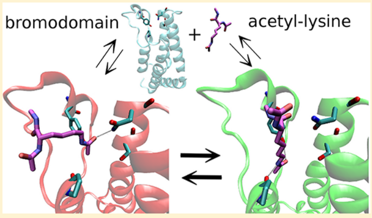
JCTC Journal of Chemical Theory and Computation
Article
pubs.acs.org/JCTC

Mechanism and Kinetics of Acetyl-Lysine Binding to Bromodomains

A. Magno,[†,‡] S. Steiner,[†,‡] and A. Caflisch*,[†]

[†]Department of Biochemistry, University of Zurich, Winterthurerstrasse 190 CH-8057 Zurich, Switzerland

Supporting Information



ABSTRACT: Bromodomains are four-helix bundle proteins that specifically recognize acetylation of lysine side chains on histones. The available X-ray structures of bromodomain/histone tail complexes show that the conserved Asn residue in the loop between helices B and C is involved in a hydrogen bond with the acetyl-lysine side chain. Here we analyze the spontaneous binding of acetyl-lysine to the bromodomain TAF1(2) by the first molecular dynamics simulations of histone mark binding to an epigenetic reader protein. Multiple events of reversible association sampled along the unbiased simulations allow us to determine the pathway and kinetics of binding. The simulations show that acetyl-lysine has two major binding modes in TAF1(2) one of which corresponds to the available crystal structures and is stabilized by a hydrogen bond to the conserved Asn side chain. The other major binding mode is more buried than in the crystal structures and is stabilized by two hydrogen bonds with conserved residues of the loop between helices Z and A. In the more buried binding conformation, three of the six structured water molecules at the bottom of the binding pocket are displaced by the acetyl-lysine side chain. The kinetic analysis shows that the two binding modes interconvert on a faster time scale with respect to the association/dissociation process. The atomic-level description of the binding pathway and binding modes is useful for the design of small molecule modulators of histone binding to bromodomains.

INTRODUCTION

Bromodomains are protein modules of about 110 residues which recognize acetylated lysine side chains mainly in histones and are thus involved in transcriptional regulation.[1,2] In the human genome 46 proteins with a total of 61 different bromodomains have been identified, with up to 6 bromodomains per protein.[3] Bromodomains occur often in tandem with a PHD module or with other bromodomain(s), suggesting that combinatorial readout may be at work in translating the histone code.[2,4,5] The potential role of bromodomains in tumors and inflammation[4,6,7] has spurred large-scale structural studies with the ultimate goal to facilitate the discovery of small-molecule inhibitors able to interfere in the process of reading acetylated lysine.[8,9] In 1999, the first three-dimensional structure of a human bromodomain was solved;[10] since then, the crystal and/or solution structures of more than 40 human bromodomains have been reported.[11,12]

All available structures show a conserved four-helix bundle topology in which the ZA-loop and BC-loop connect the first two α helices (called Z and A) and last two α helices (called B and C), respectively.[12,13] The acetyl-lysine binding site is very similar in all structures of bromodomains.[12] There are several crystal structures with bromodomains in complex with acetyl-lysine[4] (e.g., PDB codes 3P1C, 3O34, and 3O35) or peptides with two acetylated lysine side chains (PDB codes 3UVW, 3UVX, 3UVY, and 3UW9). These structures share a common intermolecular hydrogen bond between the side chain of the conserved Asn in the BC loop and the acetyl oxygen of the ligand. The binding mode is further stabilized by a network of water-mediated hydrogen bonds at the bottom of the binding groove and hydrophobic contacts.[14] In addition to the structural characteristics, thermodynamic parameters of histone tail binding have been measured by isothermal titration calorimetry using mono- and diacetylated histone peptide segments.[4] The measured dissociation constants are in the high micromolar range, and seem rather insensitive to the peptide length. In general, the dynamical character of the process of ligand binding makes it a challenging investigation subject.[15] Molecular dynamics (MD) simulations with explicit solvent provide a useful tool to study binding of small molecules to proteins at the atomic level of detail.[15−22]

Here, we have employed MD simulations in conjunction with cut-based free energy profile analysis[15,23,24] to shed light on the spontaneous association of acetyl-lysine to the second bromodomain of the human transcription initiation factor TFIID subunit 1 (abbreviated as TAF1(2)). This bromodomain was chosen based on a previous study[25] in which the side chain of the conserved Asn showed two interconverting orientations. In particular, the side chain amide nitrogen of TAF1(2) Asn1604 populates a solvent exposed orientation which is rotated by about 180 degrees with respect to the orientation observed in most crystal structures. In this context there are several interesting questions which can be addressed by simulations: Is conventional MD able to reproduce the experimentally observed binding mode of the acetyl-lysine side chain? On which time scale does spontaneous binding take place? Which role does the side chain orientation of the conserved Asn play in the binding process? Are there alternative binding modes? The simulations describe with

Received: May 3, 2013
Published: July 10, 2013

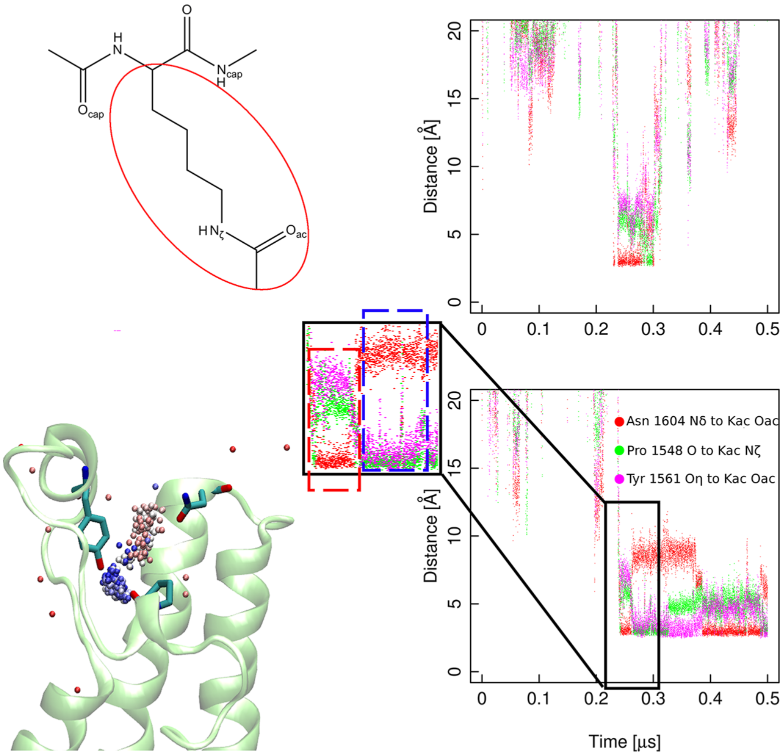


Figure 1. Spontaneous binding of Kac to TAF1(2). (top left) Two-dimensional structure of Kac. The red oval highlights the Kac side chain. (bottom left) Spheres representing the position of the carbonyl carbon atom of the Kac side chain. Their color reflects the simulation time from red to blue. The backbone of TAF1(2) and the side chains involved in binding are shown by a green ribbon and sticks, respectively. (right) Time series of key interactions along two independent simulations. The inset emphasizes the binding event shown in the bottom left panel and in Supporting Information Movie S1. The red and blue dashed rectangles in the inset illustrate the two most populated binding modes and correspond to the colors of the spheres in the bottom left panel.

atomistic resolution acetyl-lysine binding to the position observed in the crystal structure. Moreover, they reveal an alternative binding mode and provide kinetic information on the binding/unbinding processes.

■ RESULTS AND DISCUSSION

Twenty-four independent simulations of 0.5 μs each were started with acetylated lysine (with neutral blocking groups at its termini, abbreviated as Kac, see Figure 1 for the 2D structure) positioned randomly at a distance of more than 20 Å from the binding site of TAF1(2). No artificial guiding force or bias toward the binding site was used. In 16 of the 24 MD runs at least one event of spontaneous binding to the acetyl-lysine pocket was observed. In the remaining 8 MD runs Kac did not reach the binding site and remained most of the time in the solvent or was sporadically involved in unspecific association to different regions of the TAF1(2) surface. Unless stated otherwise, the following analysis refers to the 8-μs sampling of the 16 runs in which binding events were observed (a total of 4×10^5 snapshots).

Spontaneous Binding. The time series of some of the key interactions show that spontaneous and reversible binding of Kac to the acetyl-lysine pocket can take place in the 0.5 μs time scale of the MD runs (Figures 1 and 2 and Supporting Information Figures S1, S2, and movie files, ct400361k_si_001.mpg (Movie S1) and ct400361k_si_002.mpg (Movie S2)). From the analysis of hydrogen bond frequencies, it emerges that the hydrogen bond between the acetyl oxygen of the Kac side chain and the side chain of the conserved Asn1604 in the BC loop is present in nearly 40% of the snapshots of the bound state. Interestingly,

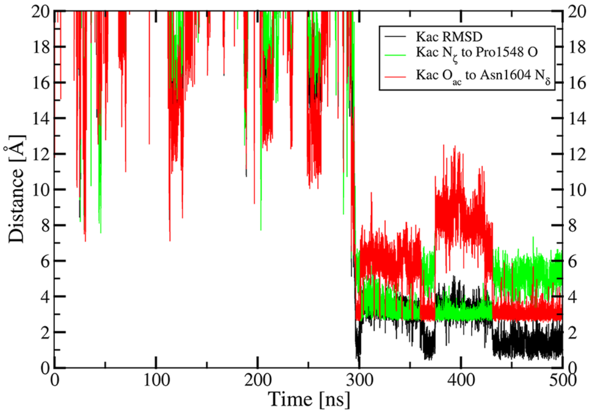


Figure 2. Spontaneous binding to the crystallographic binding mode. The plot shows the time series of the RMSD from the X-ray structure (black) together with the distances that characterize the hydrogen bond observed in the crystal structures (red) and the hydrogen bond of the more buried binding mode revealed by the present MD study (green). Note the frequent interconversions between the two binding modes. The RMSD of the Kac side chain atoms was calculated upon optimal overlap of the C_α atoms in the four helices. The reference structure for the RMSD calculation is the crystal structure of CREBBP in complex with Kac (PDB code 3P1C) as there is no crystal structure of TAF1(2) with Kac or a histone tail peptide. The MD run used for this plot is shown in Supporting Information Movie S2.

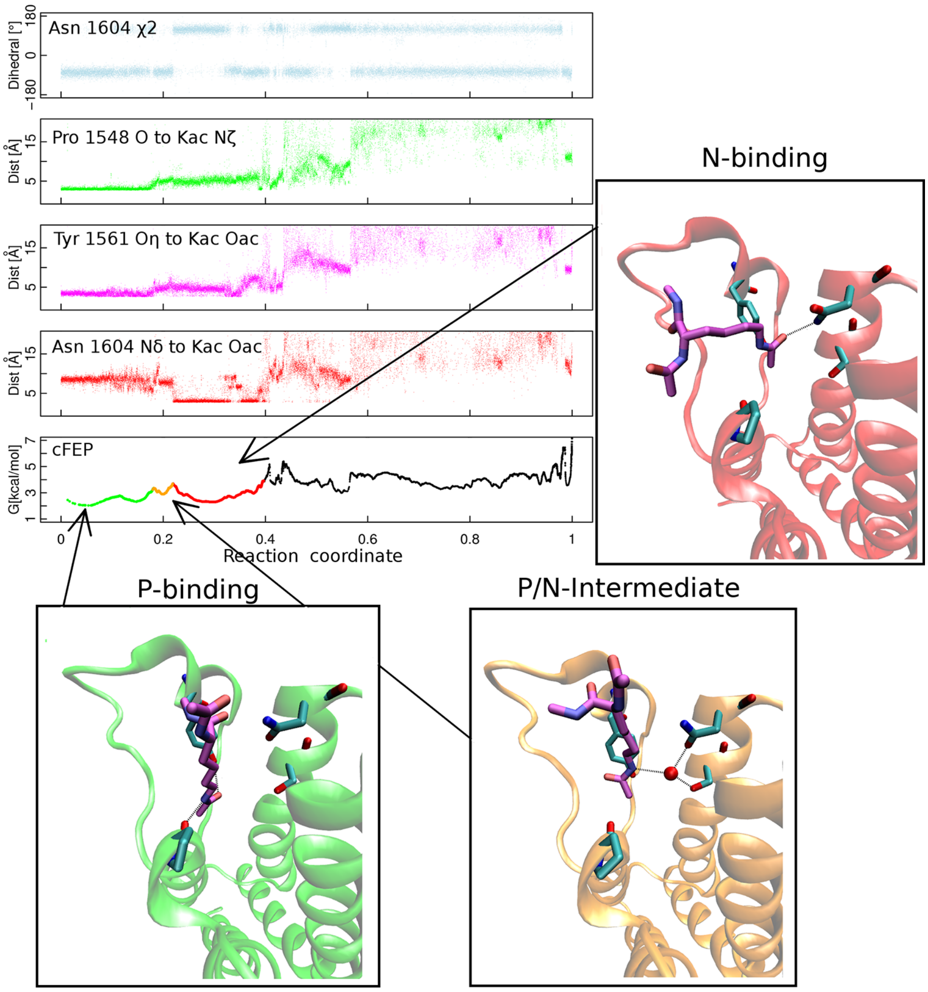


Figure 3. Geometric variables used to analyze binding (top) and cut-based free energy profile (abbreviated as cFEP) calculated using the mean first passage time from the representative node of the P-binding mode (bottom). The reaction coordinate of the cFEP is the relative partition function (see the Methods section). The colors of the cFEP reflect the three main basins whose representatives are shown with TAF1(2) in ribbon model and Kac in sticks (insets). The unbound state is in black in the cFEP. The 400 000 MD snapshots were ordered according to the free-energy sub-basin they belong to. In this way, each of the sub-basins on the free energy profile is annotated with the corresponding values of the geometric variables. As an example, the first sub-basin on the left is the P-binding mode which has hydrogen bond distance for Pro 1548 O—Kac N_ζ and Tyr1561 O_η—Kac O_{ac} whereas the Asn1604 N_δ—Kac O_{ac} distance is about 8−9 Å.

two residues of the ZA loop are also involved in hydrogen bonds with Kac. The backbone oxygen of Pro1548 (in the so-called WPF shelf) and the side chain hydroxyl of Tyr1561 act frequently as hydrogen bond acceptor and donor, respectively.

Two Major Binding Modes. The trajectories with spontaneous events of association and the corresponding time series of key intermolecular distances indicate that there are two distinct binding modes (red and blue spheres, respectively, in Figure 1 bottom, left). From the unbound state, Kac adopts first a binding mode matching the one observed in several crystal structures with the hydrogen bond between the Asn1604 N_δ atom and the Kac side chain oxygen O_{ac} (called N-binding mode hereafter). The time series of the Kac root-mean-square deviation (RMSD) from the crystallographic binding mode shows several simulation segments during which the RMSD oscillates between 0.5 and 1.5 Å (e.g., the last 60 ns of the MD run shown in Figure 2) which confirms that the N-binding mode corresponds to the binding mode determined experimentally. Note that the structure of TAF1(2) does not change upon Kac binding as monitored by the RMSD of the C_α atoms from the structure at the beginning of the MD run (see Supporting Information Figure S3) and also shown in Supporting Information Movie S2.

From the N-binding mode, Kac can reach reversibly a binding conformation which is more buried than in the crystal structure and in which the oxygen atom of Pro1548 acts as hydrogen bond acceptor for the Kac side chain N_ζ (called P-binding mode hereafter). Further stabilization is provided by a hydrogen bond between the hydroxyl group of Tyr1561 and the Kac side chain O_{ac}, as well as between atom N_δ of Asn1604 and either of the two oxygen atoms in the backbone moiety of Kac, i.e., O_{cap} (N-terminal capping group) or O (C-terminal group). An MD run with a spontaneous binding event and further interconversions between N- and P-binding modes is shown in Figure 1 and Supporting Information Movie S2.

Clustering of the snapshots based on pairwise RMSD of the Kac side chain atoms and consecutive calculation of the cut-based free energy profile[23] confirm the existence of two major binding modes (Figure 3). It is important to note that the P-binding mode is more buried (Figure 1 bottom, left) and slightly more populated than the N-binding mode. The former includes about 50% of the MD snapshots in the bound state while the latter includes about 40%. Since the P-binding mode has not been reported as of today, we decided to run additional simulations of Kac binding to TAF1(2) with a force field different from CHARMM27.[26] The N- and P-binding modes are also observed in each of four 1 μs runs with the AMBER[27] force field, and they show similar statistical weight (Supporting Information Figure S4).

We also wanted to exclude the idea that the more buried P-binding mode is an artifact due to the simplified model of the ligand. For this purpose, we run eight additional simulations of 1 μs each with the Kac-Gly-Gly-Kac tetrapeptide (derived from the diacetylated histone tail H4K5/K8) using two different bromodomains (i.e., CREBBP and BRD4(1)) and with two different force fields (CHARMM27[26] and AMBER,[27] Supporting Information Table S1). The P-binding mode is populated also by the tetrapeptide ligand in each of the two bromodomains and irrespective of the force field used (Supporting Information Figure S4). Thus, the results of the control simulations provide evidence that the P-binding mode is not an artifact of the force field or the simplified ligand. It is also important to note that in the N-binding mode the RMSD from the crystal structure of the side chain of the N-terminal Kac is almost always below 2 Å (Supporting Information Figure S5). Moreover, the C-terminal Kac, i.e., the acetyl-lysine side chain that is not in contact with the conserved Asn, has a more pronounced flexibility than the N-terminal Kac (Supporting Information Figure S5) which is consistent with the larger temperature factors of the C-terminal than the N-terminal acetyl-lysine in the crystal structure of the complex with BRD4(1) (PDB code 3UVW).

Kinetics. The existence of two major binding modes calls for further investigations of the time scales required for binding/unbinding and how they relate to the transitions within the bound state. For this analysis, first individual free energy basins were isolated by the cut-based free energy profile method following an iterative procedure reported previously[24] (see also the Methods section). Besides the two major binding modes, there is a metastable state (called P/N-intermediate and populated at about 10%) formed by MD snapshots in which the side chain of Kac points inside the binding groove but is arranged such that direct hydrogen bonds with Pro1548, Tyr1561, or Asn1604 are not possible. This binding conformation is stabilized mainly by a water-bridged hydrogen bond between Kac N_ζ and Asn1604 O_δ (see bottom, right inset in Figure 3).

The mean first passage time (MFPT) between free energy basins was evaluated using representative clusters selected according to statistical weight as published previously.[24] The MFPT values indicate that there are two well-separated time scales (Table 1). While the two major binding modes interconvert relatively fast, i.e., with an MFPT of ≈20 ns, binding/unbinding takes place on a much slower time scale (MFPT ≈ 300 ns). Figure 4 illustrates these two kinetic regimes by the free energy profiles plotted together with the MFPT to the reference state. The free energy barrier to exit the bound state (located at a value of the reaction coordinate of about 0.4 in Figure 4) is significantly higher than the barriers between the three basins in the bound state. The number of direct transitions between basins is shown in Table 1. The large number of direct transitions between P-binding and N-binding modes compared to the much smaller number of binding/unbinding events reflects the kinetic separation of the two processes which emerges also from the MFPT calculations. It is important to note that the number of binding and unbinding events (20 and 10, respectively) determined using the definition of basins by the cut-based free energy profile agrees with the visual inspection of the trajectories and the time series presented in Supporting Information Figures S1 and S2. Such agreement provides strong evidence that the clustering and the definition of free energy basins by the cut-based free energy profile method capture the relevant features of the binding and unbinding processes.

Table 1. Kinetics Data Extracted from the Simulations[a]

transitions (MFPT [ns]) from\to	unbound	P-binding	N-binding	P/N-intermediate
unbound		1 (220)	14 (220)	5 (220)
P-binding	2 (320)		129 (28)	310 (20)
N-binding	4 (320)	135 (15)		17 (29)
P/N-intermediate	4 (320)	312 (6)	16 (30)	

[a] The table shows the number of direct transitions between pairs of states as observed in the MD simulations. The MFPT values in nanoseconds are given in parentheses. Note that the MFPT values are calculated from the transition matrix, i.e., they take into account both direct and indirect transitions.

Dissociation Constant. The multiple events of spontaneous and reversible binding allow one to calculate the dissociation constant from the ratio of the off-rate and on-rate. The off-rate is the inverse of the MFPT from any of the binding modes to the unbound state (i.e., 1/320 ns^{-1}; see Table 1). The on-rate is the inverse of the MFPT to any of the bound states multiplied by the concentration of Kac in the simulation box (1/(220 ns × 6 mM)). It thus follows that the dissociation constant is about $(220 \times 6)/320 \simeq 4$ mM. This value is about 1 order of magnitude higher than the dissociation constants of mono- and diacetylated histone peptides as measured by isothermal titration calorimetry[4] which is probably a consequence of the smaller ligand used in the simulations than in the experiments.

Orientation of the Side Chain Amide Group of the Conserved Asn. The preferred orientation of the side chain of the conserved Asn1604 differs from the one observed in most bromodomain crystal structures, i.e., the N_δ atom points preferentially toward the solvent ($\chi_2 \approx -70°$, Figure 3 top) rather than the bottom of the binding site ($\chi_2 \approx 120°$). The solvent exposed orientation of the Asn1604 $-NH_2$ group is stabilized by a water-bridged hydrogen bond between the Asn side chain carbonyl and the side chain hydroxyl of Ser1600 in the C-terminal turn of helix B. This observation raises the question whether different binding modes of Kac prefer different orientations of the Asn1604 side chain. Here, it is important to note that based on a previous work[25] half of the simulation runs were started with the solvent exposed orientation of the Asn1604 —NH_2 group and the other half with its orientation toward the binding site (see Methods). Figure 5 shows the histograms of the Asn1604 χ_2 dihedral values computed over the snapshots belonging to the three individual binding modes as well as for the segments of the trajectories during which the ligand is unbound. Note, that for the unbound conformation the histogram is characterized by two peaks corresponding to the two different orientations with a preference for the solvent exposed orientation. Importantly,

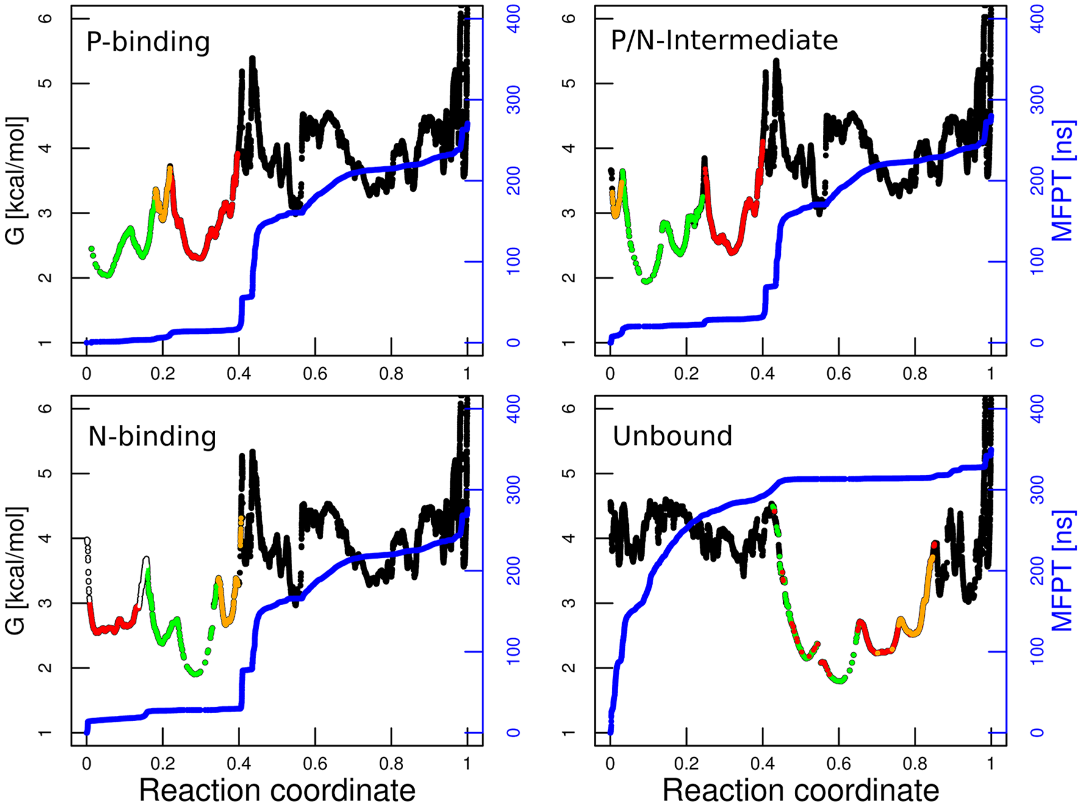


Figure 4. Four panels show the cut-based free energy profiles with the three equilibrium bound conformations or the unbound state as reference. Each panel is labeled with the state used as reference which is the first basin on the left. Individual basins are colored according to the same color scheme as in Figure 3, i.e., green, red, orange, and black for P-binding mode, N-binding mode, P/N-intermediate, and unbound state, respectively. The empty circles in the bottom left panel represent the nodes on the barrier between the N-basin and the unbound state. The reaction coordinate is the relative partition function (see the Methods section). The MFPT values to the representative node of the reference basin are also shown (blue lines with *y*-axis on the right). The MFPT is a monotonous growing function because the nodes are sorted according to their kinetic distance from the reference basin. Note that the three binding modes have approximately equal kinetic separation from the unbound conformation (i.e., 320 ns; see Table 1). Thus, there is a significant overlap of the three binding modes when plotting the profile from a node belonging to the unbound conformation (bottom right panel).

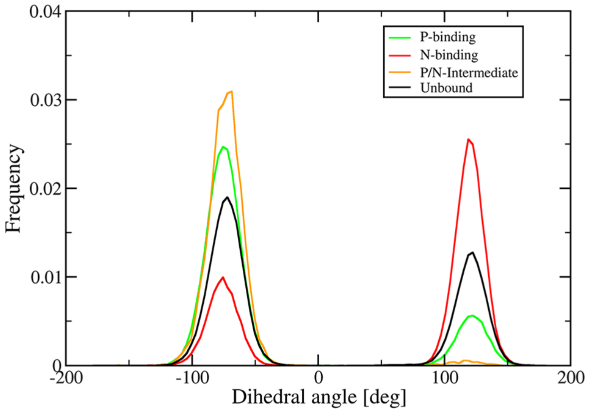


Figure 5. Histograms of the two main conformations of the χ_2 angle of the conserved Asn1604 for the individual binding modes (colors) and the unbound state (black). Positive values of the χ_2 angle correspond to the orientation of Asn1604 N_δ toward the bottom of the binding site while negative values to the orientation of Asn1604 N_δ toward the solvent.

as experimentally observed in several X-ray structures of the acetyl-lysine−bromodomain complex, the N-binding conformation peaks at $\chi_2 \approx 120°$, i.e., the orientation of Asn1604 —NH_2 pointing toward the binding site. On the contrary, in the P-binding and the P/N-intermediate modes the orientation of the —NH_2 group of Asn1604 toward the solvent is preferred, since in the former it frequently acts as donor for the O or O_{cap} atoms of Kac, while in the latter a water molecule coordinates O_δ of Asn1604, O_γ of Ser1600, and N_ζ of Kac (see Figure 3). Furthermore, it is interesting to analyze the role of Asn1604 during the binding process. For this purpose, the N-binding mode is decomposed into two sub-basins made up of snapshots with positive and negative values, respectively, of the Asn1604 χ_2 dihedral. Strikingly, 12 out of the 14 transitions from the unbound state to the N-binding state take place when the N_δ atom of the Asn1604 points toward the solvent (Supporting Information Table S2 and Movie S1). The Kac ligand can then further penetrate the binding groove, ending up in one of the major binding modes.

Conserved Water Molecules in the Acetyl-Lysine Binding Site. Six water molecules stabilized by hydrogen bonds with polar groups in the bottom of the acetyl-lysine binding site seem to be conserved in most crystal structures of bromodomains.[8] In the N-binding mode, the six conserved water molecules are arranged in the same fashion as observed in the crystal structures. In contrast, in the P-binding mode, water molecules 2, 3, and 4 (nomenclature of Figure 1 of ref 8) are replaced by the Kac side chain which, as mentioned above, is more buried than in the N-binding mode (Figure 6). This simulation result is useful for the design of small molecule modulators of histone binding to bromodomains. In particular, some of the conserved water molecules need not necessarily be present in high-throughput docking campaigns. Moreover, optimization of known binders could focus on the replacement of one or more of the structured water molecules by functional groups with similar hydrogen bonding properties.

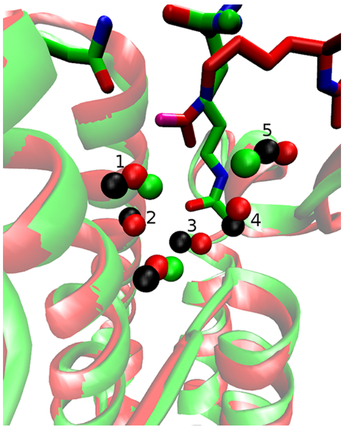


Figure 6. Structured water molecules are conserved in the N-binding mode while three of them are displaced by the tip of the Kac side chain in the P-binding mode. The representative snapshots of the N-binding mode (red) and P-binding mode (green) are aligned structurally to the crystal structure of apo TAF1(2) (PDB code 3UV4) using only the C_α atoms of the bromodomain helices. The bromodomain backbone, Kac, and structured water molecules are shown (ribbons, sticks, and spheres, respectively) for the two binding modes. For the crystal structure, only the structured water molecules are shown (black spheres) to avoid overcrowding of the figure. The Asn1604 side chain of the P-binding mode representative is also shown (carbon atoms in green). The labeling of the structured water molecules is the same as in ref 8. Water molecules 2−4 are displaced in the P-binding mode.

CONCLUSIONS

We have analyzed the pathway and kinetics of Kac binding to the TAF1(2) bromodomain by multiple MD simulations of spontaneous and reversible binding. The MD trajectories reveal two major binding modes of Kac which differ in the bromodomain residues providing the anchoring interactions. From the unbound state, Kac reaches first a position in which it is stabilized by a hydrogen bond between its side chain carbonyl oxygen and the side chain amide of the conserved Asn. This binding mode is identical to the binding mode observed in several crystal structures of bromodomains with Kac or short peptide segments from histone tails. From the crystallographic binding mode Kac further penetrates the binding site and reaches a more buried binding mode in which the Kac side chain is involved in a hydrogen bond with the carbonyl group of the conserved Pro in the so-called WPF shelf. The WPF shelf is a highly conserved motif of the bromodomain sequence and has been highlighted as an important hotspot to increase binding affinity of small molecule inhibitors.[28,29] Moreover, the binding of diacetylated histone tails involves the WPF shelf which interacts with the acetyl-lysine side chain that is not in contact with the conserved Asn.[4,30] Furthermore, the six structured water molecules observed in most crystal structures are present in the crystallographic binding conformation, while three of them are replaced by the Kac side chain in the more buried binding mode. Importantly, in several crystal structures of bromodomains in complex with inhibitors (e.g., PDB codes 4HXK, 4HXO, 4IR3, 4A9I, and 4ALC) the carbonyl oxygen of the Pro of the WPF shelf is involved in a polar interaction with the inhibitor, e.g., with an aromatic CH of the [1,2,4]triazolo-[4,3-*a*]pyridine moiety in the complex structure with BRD4(1) (PDB code 4HXO).[31] The crystal structure of the complex between BRD4(1) and 6,7-dihydrothieno[3,2-*c*]pyridin-5(4*H*)-yl(1*H*-imidazol-1-yl)methanone (PDB code 4HXK) is of particular interest as the 6,7-dihydrothieno[3,2-c]pyridin moiety is almost as buried as acetyl-lysine in the P-binding mode (Supporting Information Figure S6). Moreover, one of the structured water molecules is displaced by the dihydrothieno ring.[31]

The two binding modes observed in the MD simulations are distinct and separated by a free energy barrier which is smaller than the barrier of unbinding. In other words, upon binding there are several transitions between the two binding modes which can interconvert about 1 order of magnitude faster than the time required for full dissociation which is about 300 ns. The simulation methods used here, unbiased MD with explicit solvent (carried out on conventional compute clusters) and cut-based free energy profile analysis, can be applied to determine the binding mode(s) and kinetics of small molecules to proteins thus providing dynamical information which is useful to complement the static picture obtained by conventional experimental techniques. Although the histone code may consist of multiple post-translational modifications, the mechanism and kinetics of acetyl-lysine binding to bromodomains is a first step toward the in-depth understanding of the role of bromodomains in transcriptional regulation.

METHODS

Simulation Protocols. The coordinates of the TAF1(2) bromodomain (PDB code 3UV4) were downloaded from the protein database (URL www.pdb.org).[32] To reproduce neutral pH conditions, the side chains of aspartates and glutamates were negatively charged, those of lysines and arginines were positively charged, and the histidine side chains were neutral. Thirteen crystal water molecules inside or close to the binding site were kept while the remaining water molecules were deleted. Subsequently, the structure was solvated in a water box whose size was chosen to have a minimal distance of 12 Å between the boundary and any atom of the protein. The simulation system contained sodium and chloride ions to approximate an ionic strength of about 150 mM and to compensate for the total charge of the bromodomain. Preliminary MD simulations were carried out with CHARMM,[33] while the production runs were carried out with GROMACS 4.5.4[34] using the CHARMM PARAM27 all atom force field[26,35] and the TIP3P model of water.[36]

Periodic boundary conditions were applied, and electrostatic interactions were evaluated using the particle-mesh Ewald summation method.[37] The van der Waals interactions were truncated at a cutoff of 10 Å. The temperature of 310 K was kept constant by an external bath with velocity rescaling,[38] and the pressure was kept close to 1 atm by the Berendsen barostat.[39] The LINCS algorithm was used to fix the covalent bonds involving hydrogen atoms. The integration time step was 2 fs, and snapshots were saved every 20 ps.

Multiple independent MD runs of spontaneous binding of Kac to TAF1(2) were carried out. To mimick the context of the longer histone sequence, the N-terminus and C-terminus of Kac were neutralized by capping with acetyl and *N*-methyl-amide, respectively (Figure 1). Two different starting structures were generated for TAF1(2) by swapping the coordinates of Asn1604 O_δ and N_δ atoms in the X-ray structure (before hydrogen generation). Thus two different χ_2 dihedrals of Asn1604 were represented. Six starting positions for the Kac molecule were generated by randomly placing Kac with its center of mass ≈20 Å away from the center of mass of Asn1604 using the package PyMOL.[40] For each of the six Kac starting positions and each of the two TAF1(2) starting structures, two simulations with different seeds were started for a total 24 independent runs of 0.5 μs each (Supporting Information Table S1). Each 0.5 μs run took about 1 week on 64 cores of Xeon 5500-series processors.

Control simulations were carried out with the tetrapeptide Kac-Gly-Gly-Kac to investigate potential differences between the isolated Kac and histone tails. As in the runs with Kac, the tetrapeptide was capped with acetyl and *N*-methyl-amide groups at the N-terminus and C-terminus, respectively. The coordinates of the tetrapeptide and of the bromodomain BRD4(1) were taken from the PDB structure 3UVW. For CREBBP, the tetrapeptide was placed into its bindig site after structural overlap of the CREBBP coordinates (PDB entry code 3P1C) onto the ones of BRD4(1). From each of these two starting structures, two independent 1 μs runs were performed with the CHARMM PARAM27 force field. In addition, two independent 1 μs simulations with the AMBER force field were carried out for each of the two bromodomains. An overview of all simulations is given in Supporting Information Table S1.

Ligand Parametrization. For the simulations with the CHARMM PARAM27 force field, Kac was parametrized using the CHARMM general force field[41] for small molecules which is compatible with CHARMM PARAM27. For the simulations performed with the AMBER99 SB[27] force field, the parameters of the Kac side chains were taken from a previous computational study.[42]

Clustering and Cut-Based Free Energy Profile. The 16 MD trajectories with one or more binding events were used for further analysis (a total of 400 000 snapshots from 8 μs of MD). The snapshots were aligned structurally to the first frame of one of the MD trajectories using only the C_α atoms of the helical segments of TAF1(2). Subsequently, the snapshots were clustered using a tree-based algorithm[43] implemented in CAMPARI.[44] As a measure of pairwise similarity, the root-mean-square deviation (RMSD) of the acetyl-lysine side chain heavy atoms including C_α was computed without further alignment. The clustering was carried out with eight levels of the tree using 8 and 1 Å as RMSD thresholds for the coarsest and finest clustering level, respectively. This led to a total number of 149 024 clusters 126 552 of which contained only one snapshot (called singletons). The number of singletons seems large, but it has to be noted that in the unbound state, where the peptide moves freely around the bromodomain, the RMSD between two relative positions of Kac is almost always larger than 1 Å and thus the unbound state consists mainly of singletons. Other clustering methods which are able to partially solve the problem of the singletons were also tested but led to worse kinetic separation of the different states. Upon clustering, a free energy profile was computed using a method based on the equilibrium kinetic network which preserves the free energy barriers.[23,24,45] This approach emulates the cuts in flow-networks, and the computed profile is therefore named cut-based free energy profile. The nodes and links of the equilibrium kinetic network are the clusters (determined by the tree-based algorithm as explained above) and the direct transitions between them sampled along the MD runs, respectively.[46] For each node, the set of nodes is partitioned into two groups *A* and *B* using the mean first passage time (MFPT) to the reference node as an order parameter. The free energy is related to the maximum flow between sets *A* and *B* and calculated as $G = -kT\ \ln(Z_{AB}/Z)$, where Z_{AB}/Z is the relative partition function which represents the statistical weight of the transitions between sets *A* and *B* (see refs 23 and 24 for details). The result is a one-dimensional profile along the reaction coordinate Z_A/Z (i.e., the relative partition function representing the statistical weight of set *A*) which preserves the barrier height between the free energy basins. The evaluations of cut-based free energy profiles, interatomic distances, and dihedral angles were performed with the simulation analysis package WORDOM.[47,48]

Identification of Basins. The free energy basins were isolated from the cut-based free energy profile using a previously published procedure.[24] First the most populated cluster is used as reference to isolate the first basin which consists of the set of clusters on the left side of the first barrier on the profile. Due to possible overlap on the right side of the barrier (i.e., similar kinetic distance from two or more basins to the reference cluster), the isolation of each basin requires a separate profile[24] (Supporting Information Figures S7−S10). In other words, the procedure used for the first basin is iterated for the second basin which is isolated using as a reference cluster the largest cluster that is not in the first basin. Clusters on top of barriers are not assigned to any basin to avoid spurious fluctuations, i.e., shortcuts along the trajectories. Transitions between basins were calculated from the MD trajectories when the system passed from one basin to another either by a direct transition or via the not assigned snapshots (the barrier region).

Analysis of Hydrogen Bonds. Hydrogen bonds were analyzed using GROMACS.[34] A hydrogen bond between a donor and an acceptor atom was considered to be formed if the donor−acceptor distance was lower than 3.5 Å and the acceptor−donor−hydrogen angle lower than 30°.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Simulation protocol (Table S1); transition matrix computed after splitting of the N-binding mode according to the χ_2 dihedral angle of the conserved Asn (Table S2); time series of key interatomic distances for all the performed simulations (Figures S1−S4); time series of the root-mean-square displacement from the 3P1C crystal structure along a 500 ns simulation segment (Figure S3, bottom); time series of the root-mean-square displacement of Kac-Gly-Gly-Kac in complex with BRD4(1) and CREBBP from the 3UVW crystal structure (Figure S5); structural overlap of the representative snapshots of P- and N-binding mode with the structure of the complex between BRD4(1) and inhibitor reported in 4HXK (Figure S6); geometric annotation and cut-based free energy profile projections with representative nodes of the P-binding mode, N-binding mode, P/N-intermediate, and unbound state as reference nodes (Figure S7−S10); Movies S1 and S2 as mentioned in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: caflisch@bioc.uzh.ch. Phone: (+41 44) 635 55 21. Fax: (+41 44) 635 68 62.

Author Contributions

‡A.M. and S.S.: These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Drs. Danzhi Huang and Andreas Vitalis for interesting discussions. We also thank Armin Widmer for continuous support with the program WITNOTP. The simulations were carried out on the Schrödinger cluster at the University of Zurich. This work was supported by grants of the Swiss National Science Foundation and the Swiss Cancer League (Krebsliga).

REFERENCES

(1) Zeng, L.; Zhou, M.-M. *FEBS Lett.* **2002**, *513*, 124−128.
(2) Filippakopoulos, P.; Knapp, S. *FEBS Lett.* **2012**, *586*, 2692−2704.
(3) Arrowsmith, C. H.; Bountra, C.; Fish, P. V.; Lee, K.; Schapira, M. *Nat. Rev. Drug. Discov.* **2012**, *11*, 384−400.
(4) Filippakopoulos, P.; et al. *Nature* **2010**, *468*, 1067−1073.
(5) Jenuwein, T.; Allis, C. D. *Science* **2001**, *293*, 1074−1080.
(6) Zuber, J.; et al. *Nature* **2011**, *478*, 524−528.
(7) Dawson, M. A.; Kouzarides, T.; Huntly, B. J. *N. Engl. J. Med.* **2012**, *367*, 647−657.
(8) Hewings, D. S.; Rooney, T. P. C.; Jennings, L. E.; Hay, D. A.; Schofield, C. J.; Brennan, P. E.; Knapp, S.; Conway, S. J. *J. Med. Chem.* **2012**, *55*, 9393−9413.
(9) Furdas, S. D.; Carlino, L.; Sippl, W.; Jung, M. *Med. Chem. Commun.* **2012**, *3*, 123−134.
(10) Dhalluin, C.; Carlson, J. E.; Zeng, L.; He, C.; Aggarwal, A. K.; Zhou, M.-M. *Nature* **1999**, *399*, 491−496.
(11) Mujtaba, S.; Zeng, L.; Zhou, M. *Oncogene* **2007**, *26*, 5521−5527.
(12) Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J.-P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; Gingras, A.-C.; Arrowsmith, C. H.; Knapp, S. *Cell* **2012**, *149*, 214−231.
(13) Vidler, L. R.; Brown, N.; Knapp, S.; Hoelder, S. *J. Med. Chem.* **2012**, *55*, 7346−7359.
(14) Owen, D. J.; Ornaghi, P.; Yang, J.-C.; Lowe, N.; Evans, P. R.; Ballario, P.; Neuhaus, D.; Fileticі, P.; Travers, A. A. *EMBO J.* **2000**, *19*, 6141−6149.
(15) Huang, D.; Caflisch, A. *PLoS. Comput. Biol.* **2011**, *7*, e1002002.
(16) Seco, J.; Luque, F. J.; Barril, X. *J. Med. Chem.* **2009**, *52*, 2363−71.
(17) Guvench, O.; MacKerell, A. D. J. *PLoS. Comput. Biol.* **2009**, *5*, e1000435.
(18) Basse, N.; Kaar, J. L.; Settanni, G.; Joerger, A. C.; Rutherford, T. J.; Fersht, A. R. *Chem. Biol.* **2010**, *17*, 46−56.
(19) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. *J. Am. Chem. Soc.* **2009**, *131*, 11811−11818.
(20) Buch, I.; Giorgino, T.; De Fabritiis, G. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10184−10189.
(21) Huang, D.; Caflisch, A. *ChemMedChem* **2011**, *6*, 1578−1580.
(22) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13118−13123.
(23) Krivov, S. V.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 12689−12698.
(24) Krivov, S. V.; Muff, S.; Caflisch, A.; Karplus, M. *J. Phys. Chem. B* **2008**, *112*, 8701−8714.
(25) Steiner, S.; Magno, A.; Huang, D.; Caflisch, A. *FEBS Lett.* **2013**, *587*, 2158−2163.
(26) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(27) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(28) Chung, C. W.; et al. *J. Med. Chem.* **2011**, *54*, 3827−3838.
(29) Chung, C. W.; Dean, A. W.; Woolven, J. M.; Bamborough, P. *J. Med. Chem.* **2012**, *55*, 576−586.
(30) Moriniere, J.; Rousseaux, S.; Steuerwald, U.; Soler-Lopez, M.; Curtet, S.; Vitte, A.-L.; Govin, J.; Gaucher, J.; Sadoul, K.; Hart, D. J.; Krijgsveld, J.; Khochbin, S.; Muller, C. W.; Petosa, C. *Nature* **2009**, *461*, 664−668.
(31) Zhao, L.; et al. *J. Med. Chem.* **2013**, *56*, 3833−3851.
(32) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(33) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(34) Spoel, D. V. D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(35) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L. J.; J., D. E.; et al. *J. Phys. Chem. B* **1998**, *102*, 35863616.
(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926935.
(37) Darden, T.; York, D.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.
(38) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
(39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(40) *The PyMOL Molecular Graphics System*, version 1.3r1; Schrödinger, LLC, New York, 2010.
(41) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. *J. Comput. Chem.* **2010**, *31*, 671−690.
(42) Liu, H.; Duan, Y. *Biophys. J.* **2008**, *94*, 4579−4585.
(43) Vitalis, A.; Caflisch, A. *J. Chem. Theory Comput.* **2012**, *8*, 1108−1120.
(44) Vitalis, A.; Steffen, A.and Lyle, N.; Mao, A. H.; Pappu, R. V. *CAMPARI v1.0.* http://sourceforge.net/projects/campari (accessed April 22, 2013).
(45) Krivov, S. V.; Karplus, M. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13841−13846.
(46) Rao, F.; Caflisch, A. *J. Mol. Biol.* **2004**, *342*, 299−306.
(47) Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A. *Bioinformatics* **2007**, *23*, 2625−2627.
(48) Seeber, M.; Felline, A.; Raimondi, F.; Muff, S.; Friedman, R.; Rao, F.; Caflisch, A.; Fanelli, F. *J. Comput. Chem.* **2011**, *32*, 1183−1194.

Mechanism and kinetics of acetyl-lysine binding to bromodomains

SUPPLEMENTARY INFORMATION

A. Magno[a‡], S. Steiner[a‡], and A. Caflisch[a*]

[a]Department of Biochemistry
University of Zürich, Winterthurerstrasse 190
CH-8057 Zürich, Switzerland
Phone: (+41 44) 635 55 21, FAX: (+41 44) 635 68 62
email: caflisch@bioc.uzh.ch

*Corresponding author

‡ These authors contributed equally to this work

keywords: molecular dynamics, post-translational modifications, histones, epigenetic marks, acetylated lysine, cut-based free energy profile

July 4, 2013

Bromodomain (PDB code)	Ligand	Force Field	Starting structure	No. of runs	Total sampling [μs]
TAF1(2) (3UV4)	Kac	CHARMM	Unbound, Asn1604 χ_2 up	12	6
TAF1(2) (3UV4)	Kac	CHARMM	Unbound, Asn1604 χ_2 down	12	6
TAF1(2)	Kac	AMBER	P-binding from CHARMM	2	2
TAF1(2)	Kac	AMBER	N-binding from CHARMM	2	2
BRD4(1) (3UVW)	Tetrapeptide	CHARMM	X-Ray	2	2
CREBBP (3P1C)	Tetrapeptide	CHARMM	Overlap on 3UVW	2	2
BRD4(1)	Tetrapeptide	AMBER	from CHARMM run	2	2
CREBBP	Tetrapeptide	AMBER	from CHARMM run	2	2

Table S1: Simulations performed.

Transitions from \ to	**Unbound**	**P-binding**	**N-binding DOWN**	**N-binding UP**	**P/N-Intermediate**
Unbound	-	1	2	12	5
P-binding	2	-	119	10	310
N-binding DOWN	2	118	-	18	4
N-binding UP	2	17	18	-	13
P/N-Intermediate	4	312	6	10	-

Table S2: The number of transitions between the different basins after splitting of the N-binding basin into two sub-basins characterized by different values of Asn1604 χ_2.

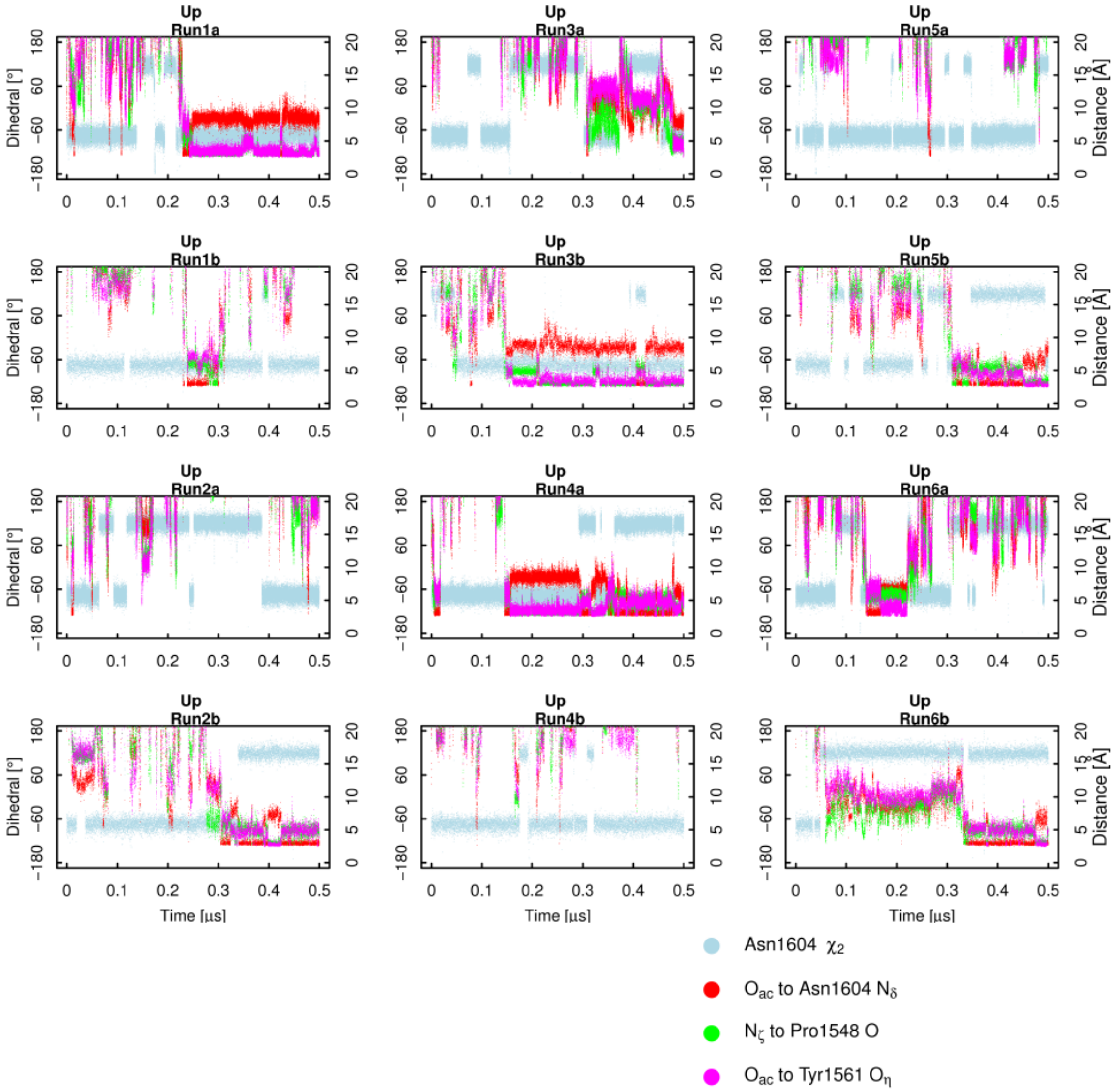


Figure S1: Time series of the interatomic distances between donor and acceptor atoms of three intermolecular key interactions (red, green, and magenta, y-axis on the right) together with the time series of the χ_2 dihedral angle of the conserved Asn1604 (cyan, y-axis on the left) for the 12 individual runs started with the N_δ atom of Asn1604 pointing towards the solvent (Up). The letters *a* and *b* distinguish two runs started from the same starting structure using different seeds.

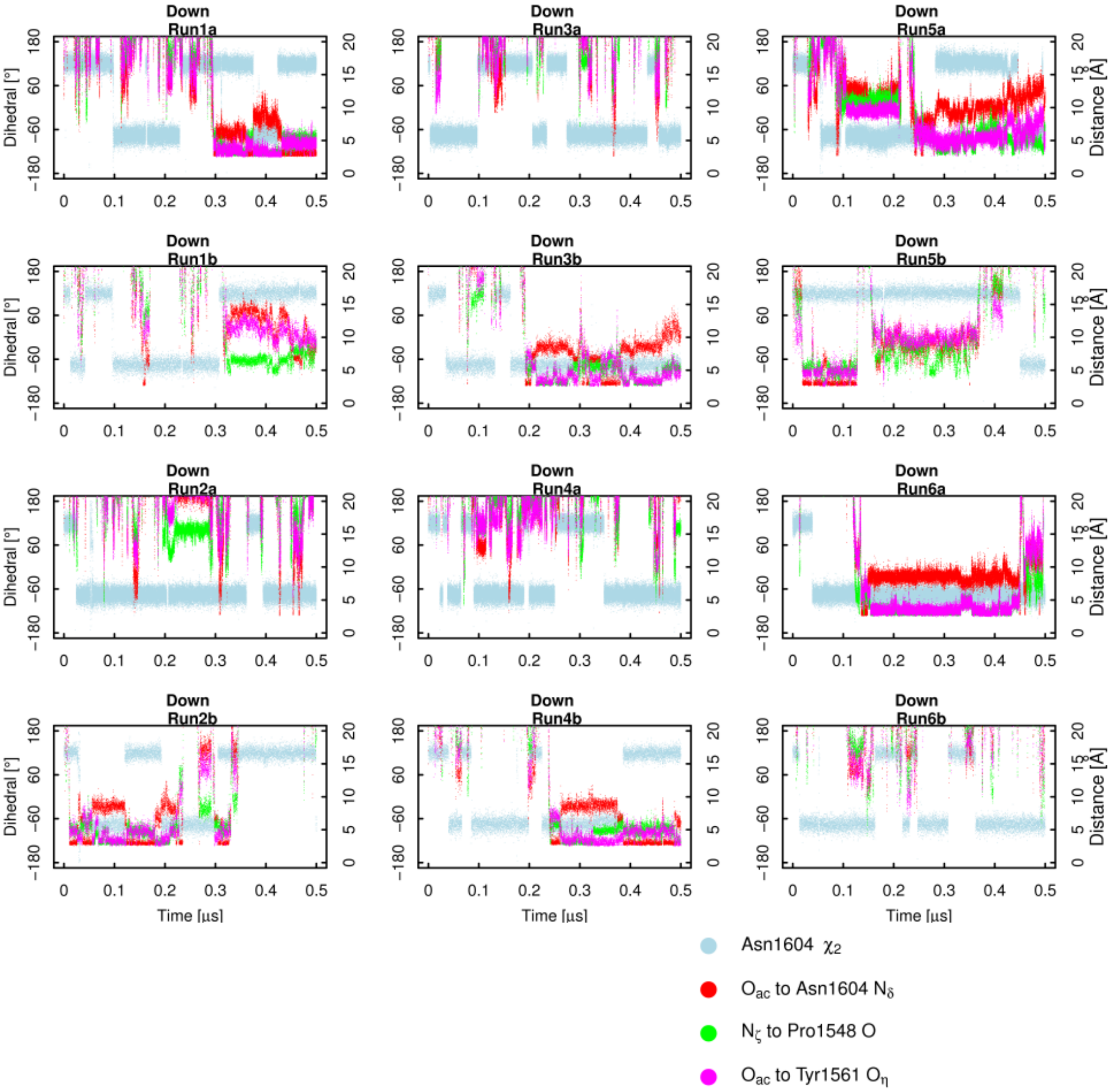


Figure S2: Same as in Figure S1 for the 12 individual runs started with the N_δ atom of Asn1604 pointing towards the binding site bottom (Down).

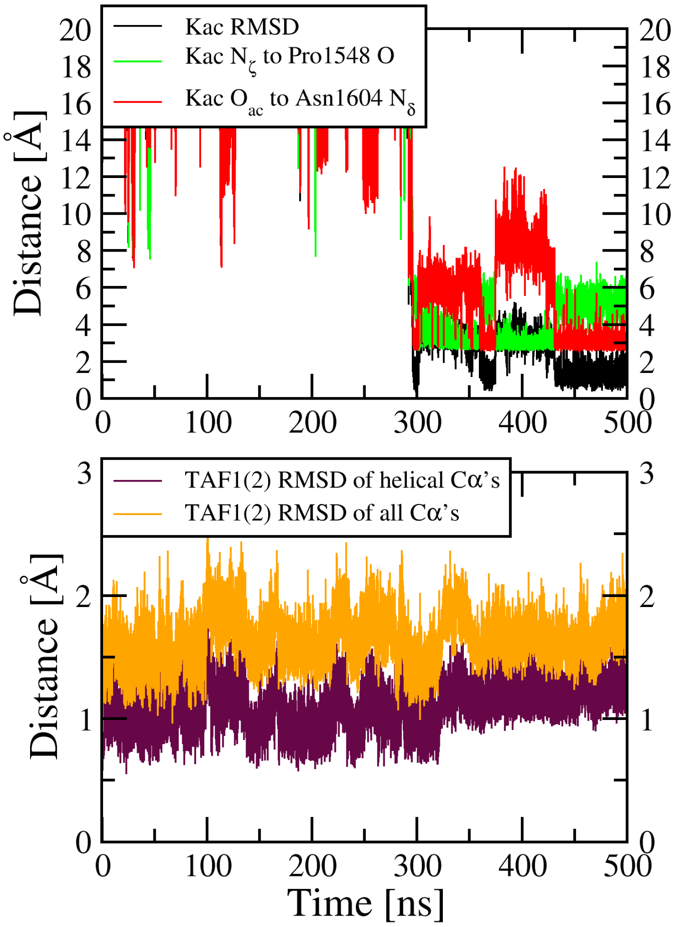


Figure S3: (Top) Time series of the RMSD from the X-ray structure (black) together with the distances that characterize the hydrogen bond observed in the crystal structures (red) and the hydrogen bond of the more buried binding mode revealed by the present MD study (green). The RMSD of the Kac side chain atoms was calculated upon optimal overlap of the C_α atoms in the four helices. The reference structure for the RMSD calculation was the crystal structure of CREBBP in complex with Kac (PDB code 3P1C) as there is no crystal structure of TAF1(2) with Kac or a histone tail peptide. (Bottom) Time series of the RMSD from the equilibrated conformation of TAF1(2) C_α atoms (orange) and C_α atoms in helices only (maroon). These time series show that upon Kac binding the bromodomain structure becomes slightly less flexible but does not undergo any relevant conformational change.

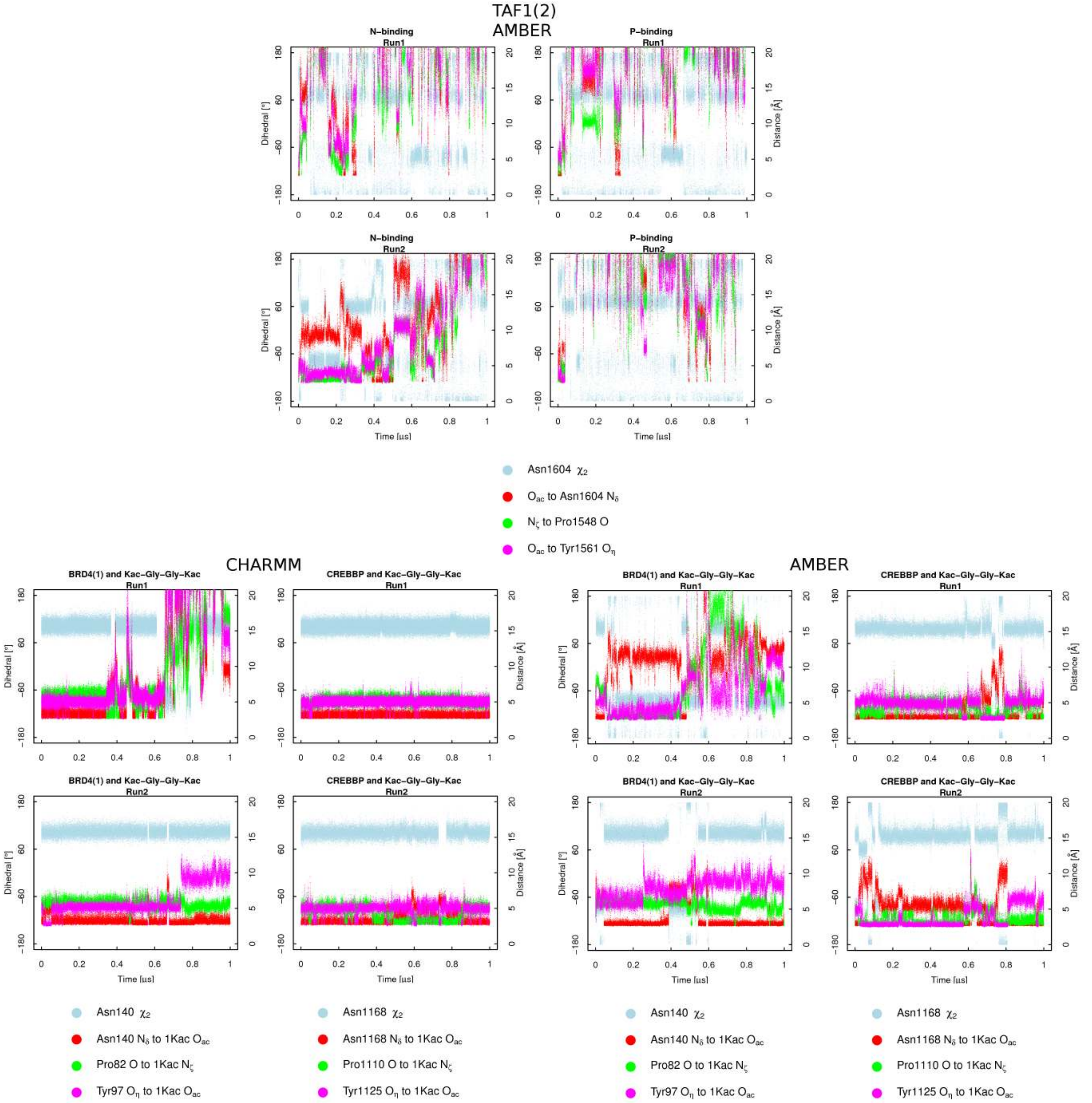


Figure S4: Control simulations with the AMBER force field and/or the tetrapeptide Kac-Gly-Gly-Kac. (Top) Time series of the interatomic distances between donor and acceptor atoms of three intermolecular key interactions (red, green, and magenta, y-axis on the right) together with the time series of the χ_2 dihedral angle of the conserved Asn in the binding site (cyan, y-axis on the left) for the control simulations of TAF1(2) and Kac with AMBER force field. (Bottom) Time series of interatomic distances between the N-terminal Kac of the tetrapeptide Kac-Gly-Gly-Kac and BRD4(1) or CREBBP for the simulations with the CHARMM (bottom left) or AMBER (bottom right) force field.

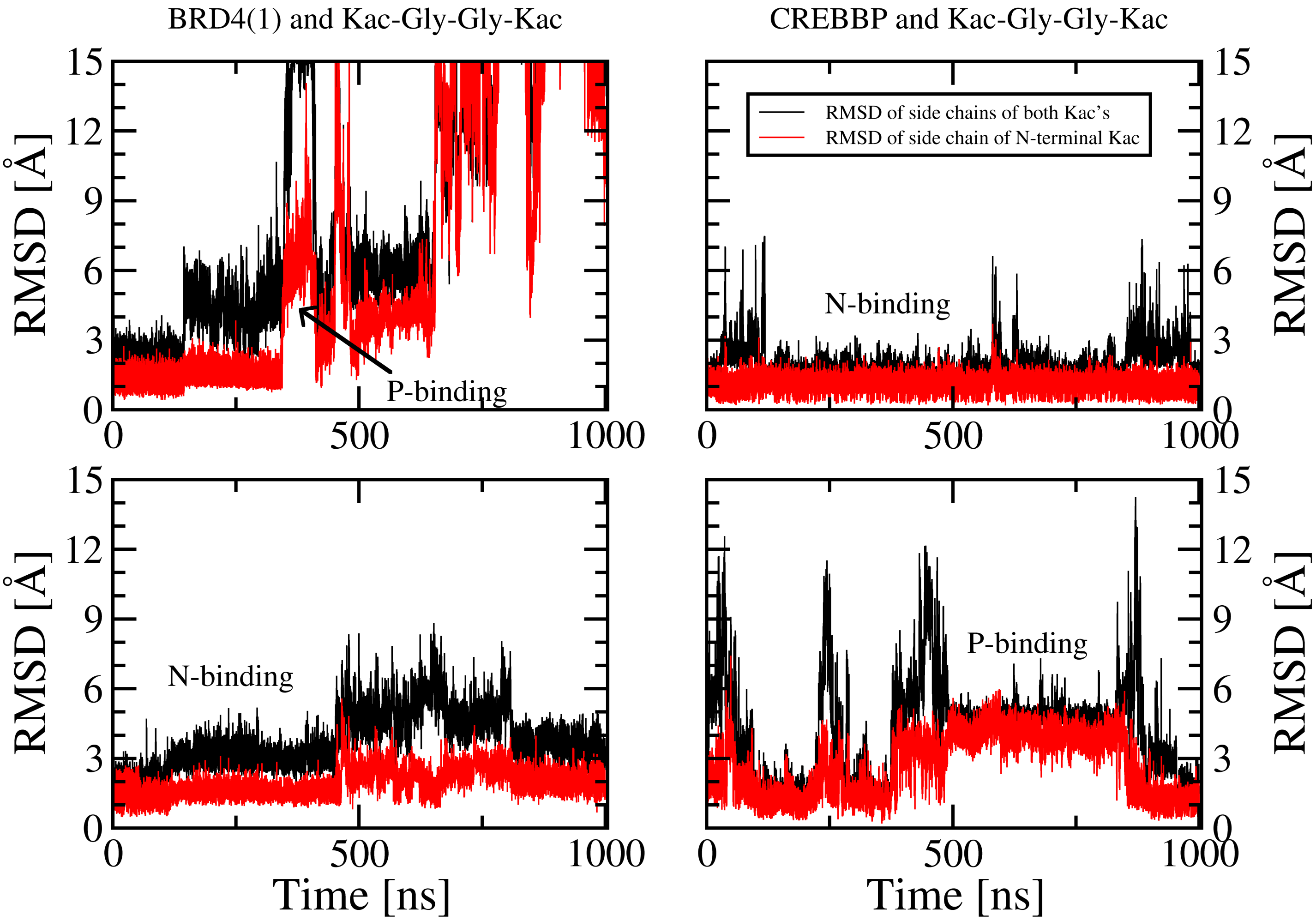


Figure S5: Control simulations with the CHARMM force field and the tetrapeptide Kac-Gly-Gly-Kac. Time series of the RMSD of the side chain atoms of the two Kac's (black) and only the N-terminal Kac (red) for two independent runs each with BRD4(1) (left) and CREBBP (right). The RMSD was calculated upon optimal overlap of the C_α atoms in the four helices of the bromodomain. The reference structure for the RMSD calculation was the crystal structure of BRD4(1) in the complex with Kac-Gly-Gly-Kac from the histone 4 peptide H4K5acK8ac (PDB code 3UVW).

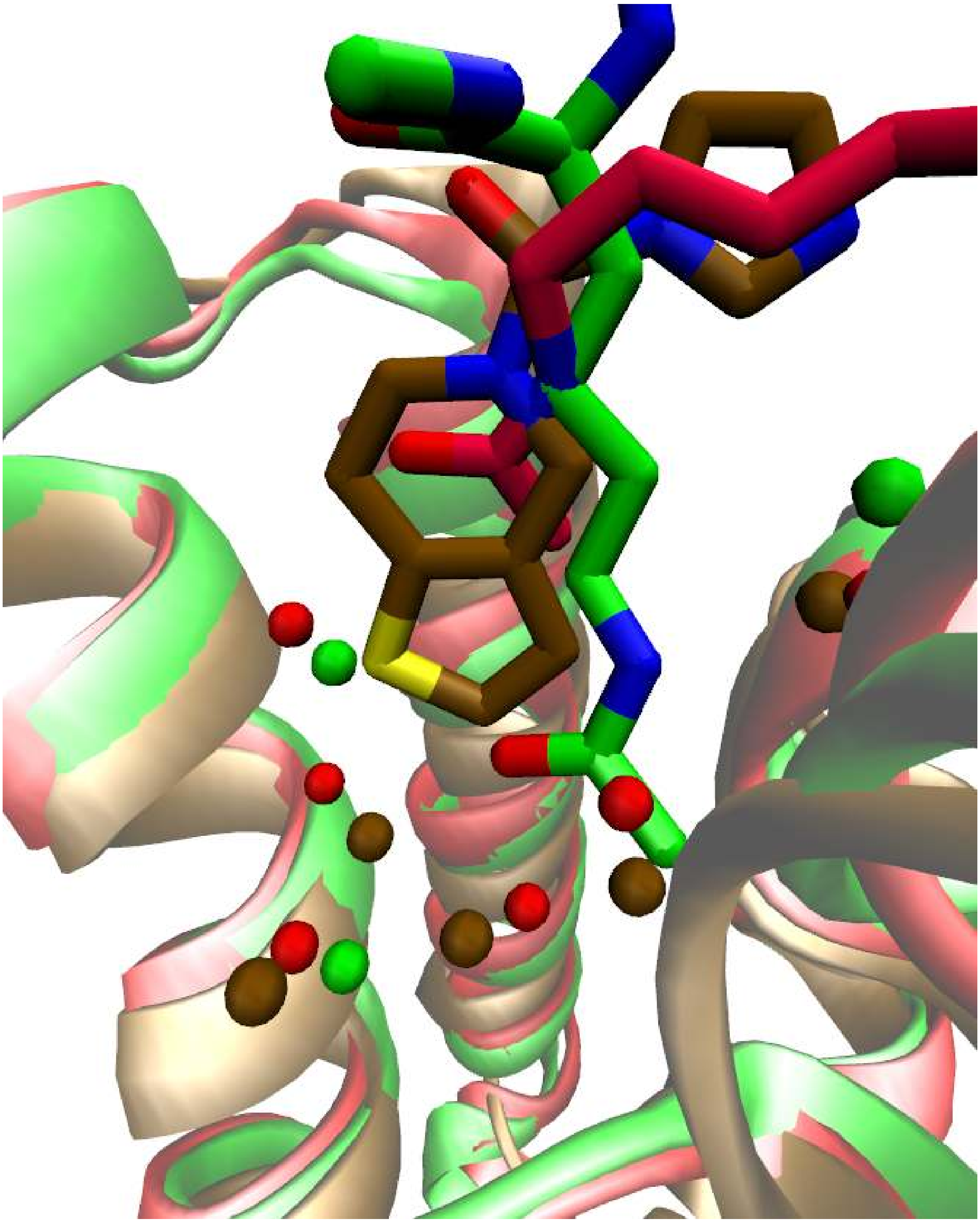


Figure S6: Small molecule inhibitors of brodomains can penetrate deeply into the acetyl-lysine binding pocket similarly to the P-binding mode of Kac observed in the MD simulations. Representative snapshots of acetyl-lysine in the P-binding mode (green) and N-binding mode (red) are structurally aligned, using only the C_α atoms of the bromodomain helices, to the crystal structure of BRD4(1) in complex with a recently published inhibitor (brown, PDB code 4HXK). The bromodomain backbone, Kac/inhibitor, and structured water molecules are shown as ribbons, sticks and spheres, respectively.

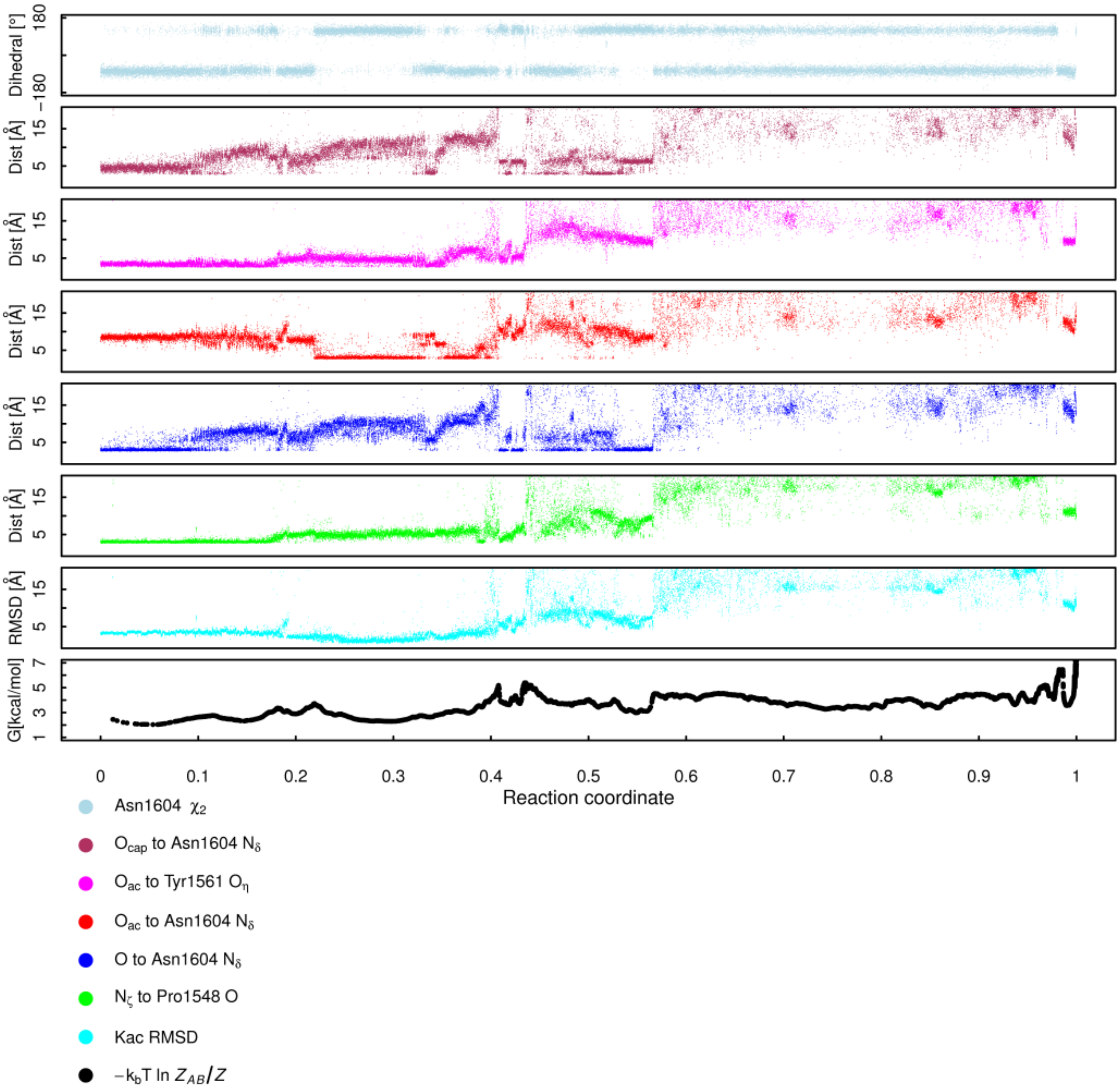


Figure S7: Geometric annotation and cFEP (bottom panel) with the P-binding mode representative as reference node. The RMSD (second panel from the bottom) of the Kac atoms with respect to the X-ray structure reported in 3P1C was calculated after overlap of the TAF1(2) C_α atoms to the corresponding atoms of CREBBP in 3P1C.

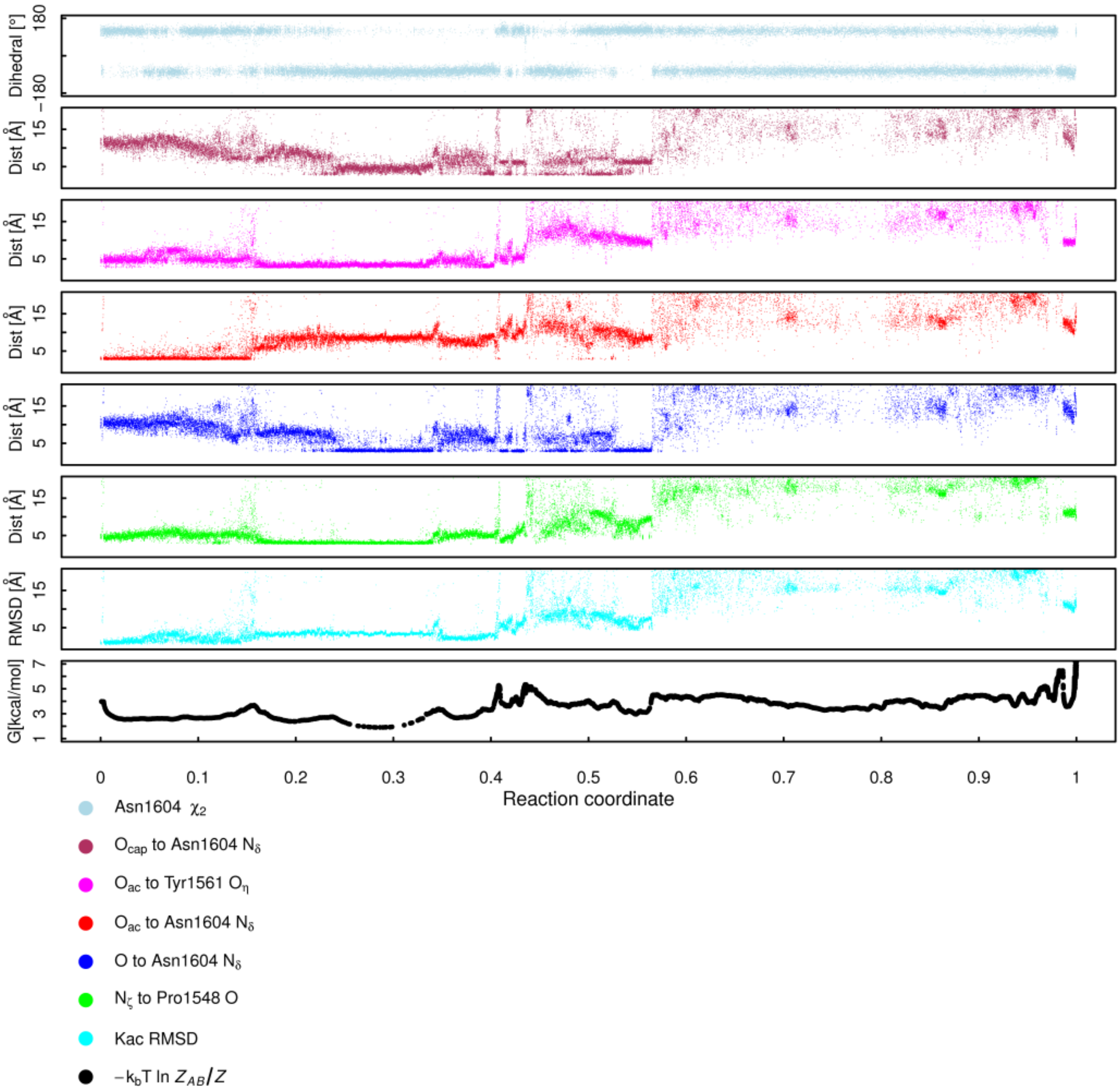


Figure S8: Geometric annotation and cFEP (bottom panel) with the N-binding mode representative as reference node. For further details see legend of Figure S7.

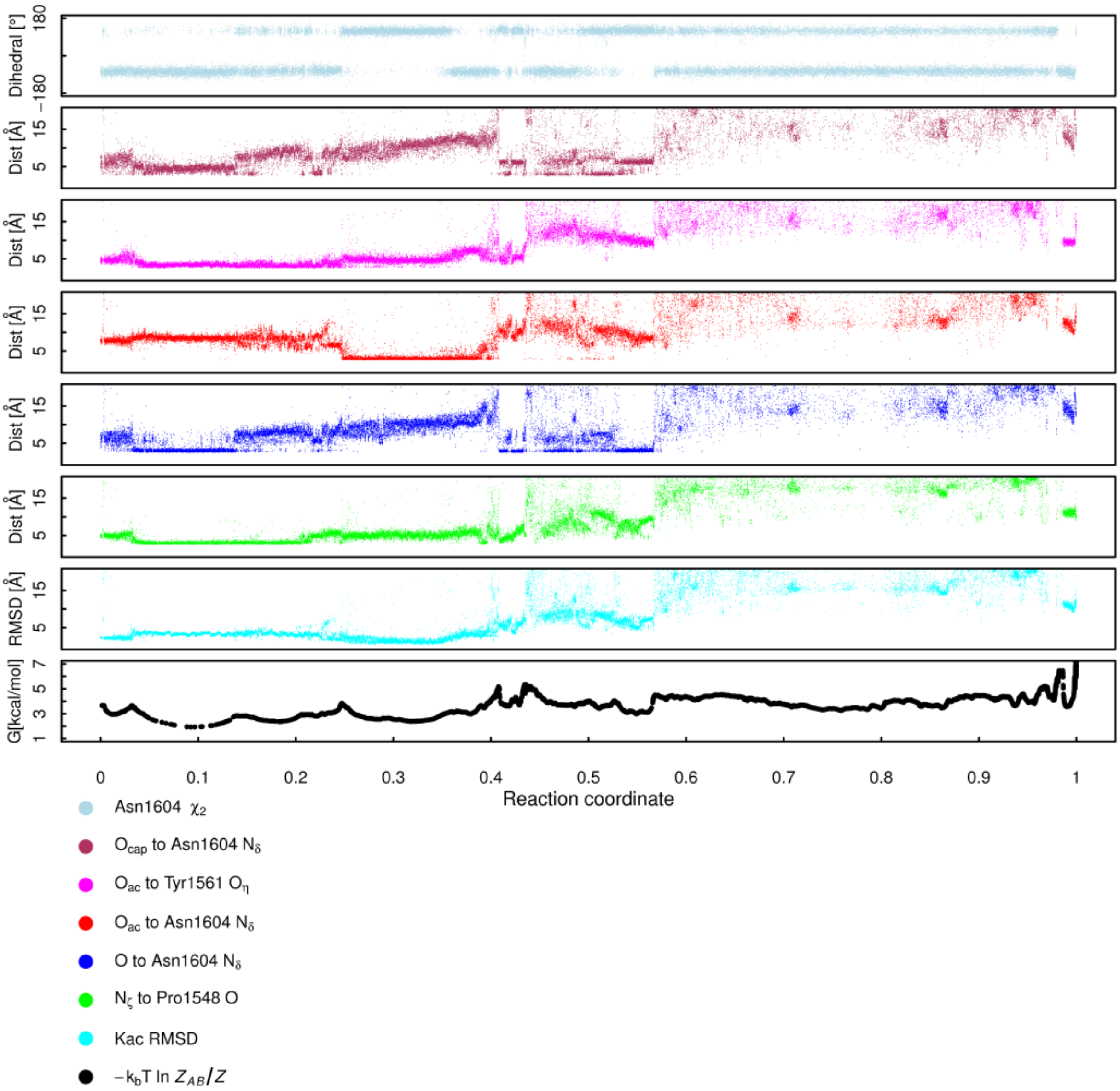


Figure S9: Geometric annotation and cFEP (bottom panel) with the P/N-Intermediate representative as reference node. For further details see legend of Figure S7.

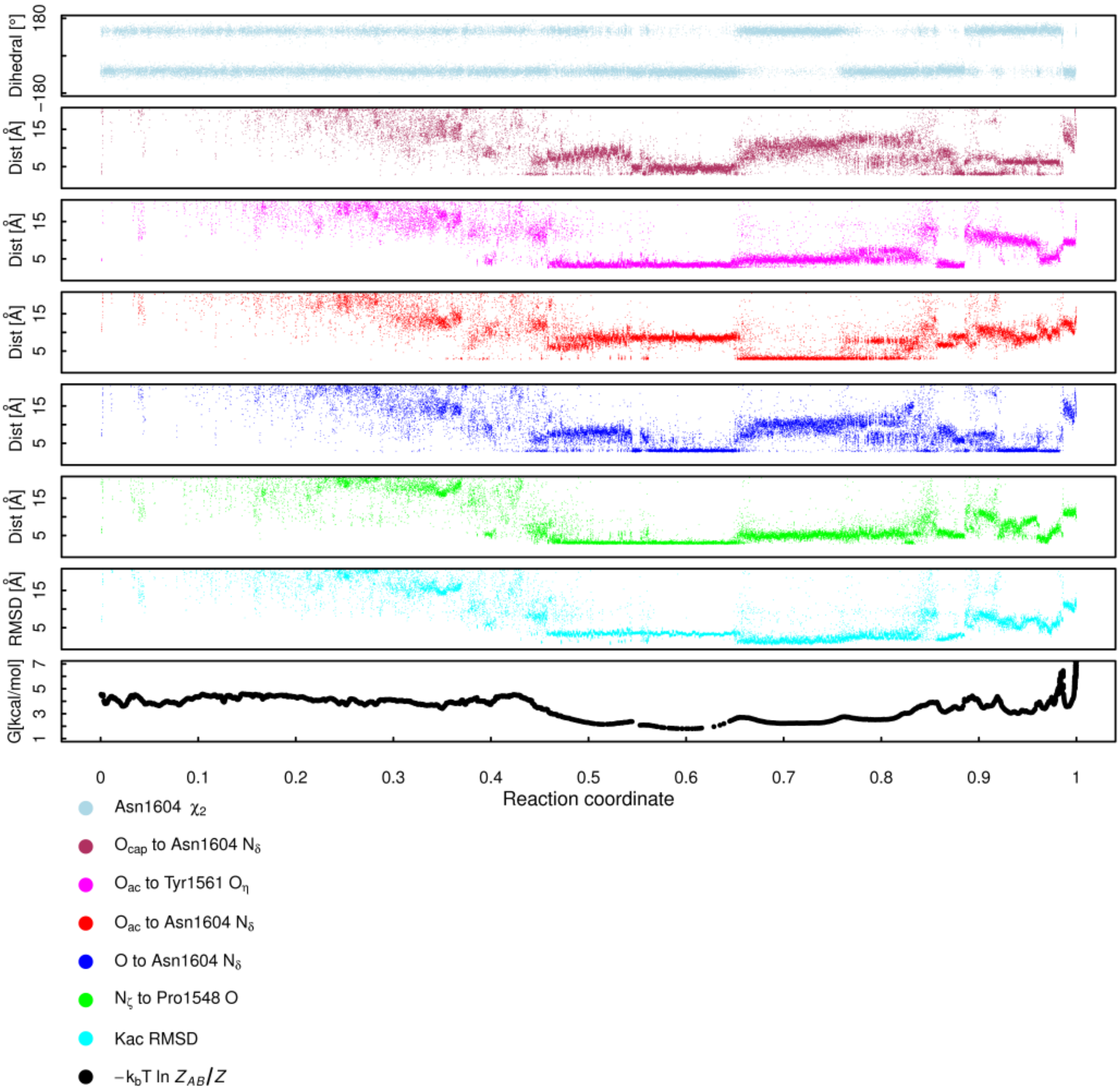


Figure S10: Geometric annotation and cFEP (bottom panel) from the unbound state. For further details see legend of Figure S7.

Chapter 7

Does bromodomain flexibility influence histone recognition?

Steiner, S., Magno, A., Huang, D. and Caflisch, A. *FEBS Lett.*, **2013**, *587, 2158-2163*

FEBS Letters 587 (2013) 2158–2163

FEBS Letters

journal homepage: www.FEBSLetters.org

Does bromodomain flexibility influence histone recognition?

Sandra Steiner, Andrea Magno, Danzhi Huang, Amedeo Caflisch *

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 27 February 2013
Revised 10 May 2013
Accepted 13 May 2013
Available online 25 May 2013

Edited by Robert B. Russell

Keywords:
Molecular dynamics
Post-translational modification
Histone
Epigenetic target

ABSTRACT

Bromodomains are protein modules that selectively recognize histones by binding to acetylated lysines. Here, we have carried out multiple molecular dynamics simulations of 20 human bromodomains to investigate the flexibility of their binding site. Some bromodomains show alternative side chain orientations of three evolutionarily conserved residues: the Asn involved in acetyl-lysine binding and two conserved aromatic residues. Furthermore, for the BAZ2B and CREBBP bromodomains we observe occlusion of the binding site which is coupled to the displacement of the two aromatic residues. In contrast to available structures, the simulations reveal large variability of the binding site accessibility. The simulations suggest that the flexibility of the bromodomain binding site and presence of self-occluded metastable states influence the recognition of acetyl-lysine on histone tails.

1. Introduction

Bromodomains, protein modules of about 110 residues, recognize acetylated lysine side chains mainly in histones and are thus involved in transcriptional regulation [1,2]. In the human genome there are 46 proteins with a total of 61 different bromodomains, with up to six bromodomains per protein [3]. Due to the potential role of bromodomains in tumors and inflammation [4–6], large scale structural studies have been carried out with the ultimate goal to facilitate the discovery of small-molecule inhibitors able to interfere in the process of reading acetylated lysine [7]. Since 1999, when the first three-dimensional structure of a human bromodomain was solved [8], the crystal and/or solution structures of more than 40 human bromodomains have been reported [9,10]. All available structures show a conserved four-helix bundle topology (Fig. 1) in which the ZA-loop and BC-loop connect the first two α helices (called Z and A) and last two α helices (called B and C), respectively [10,11]. The acetyl-lysine binding site is very similar in all structures of bromodomains [10]. The BC-loop contributes the evolutionary conserved Asn side chain [12] which acts as hydrogen bond donor to the acetylated lysine side chain. Moreover, two conserved Tyr residues [13] are present in the ZA-loop and BC-loop, respectively (Fig. S1). In striking contrast to the abundance of available three-dimensional structures, there are only few computational studies on bromodomains [14–16]. In particular, it seems that the dynamic properties of human apo bromodomains have not been investigated yet by atomistic simulations. Here, we have studied the binding site flexibility of 20 bromodomains (covering seven of the eight families of human bromodomains) by explicit solvent MD simulations. The simulation results reveal a surprisingly high heterogeneity of the plasticity and accessibility of the acetyl-lysine binding site.

2. Results and discussion

A sequence alignment of the 20 simulated bromodomains with indication of important structural features and investigated key residues is given in Fig. S1. Family memberships are listed in Table S-I. The time series of root mean square deviation (RMSD) of the C_α atoms show that the overall fold of the simulated bromodomains is stable over the 1-µs time scale and 310 K temperature of the MD runs (Fig. S2). Most of the fluctuations are localized in the ZA-loop, BC-loop, and the termini (Fig. S3). The ZA-loop and BC-loop are involved in binding the histone tail whereas the termini are located far away from the acetyl-lysine binding site so that their motion can be neglected. Importantly, we observe multiple events of reversible rotation of the side chains investigated in the present study along the 0.5- to 1-µs time scale of the MD runs (Figs. S4 and S5). These rotations take place on a time scale about one order of magnitude smaller than the length of the simulations so that the following analysis is not marred by statistical errors.

We first investigated the flexibility of the conserved Asn which is directly involved in acetyl-lysine binding (i.e., Asn1944 in BAZ2B, Asn1168 in CREBBP, Asn1604 in TAF1(2), etc.). For most of the bromodomains studied here, there are many events of reversible

---

* Corresponding author. Fax: +41 (44) 635 68 62.
E-mail address: caflisch@bioc.uzh.ch (A. Caflisch).

0014-5793/$36.00 © 2013 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.febslet.2013.05.032

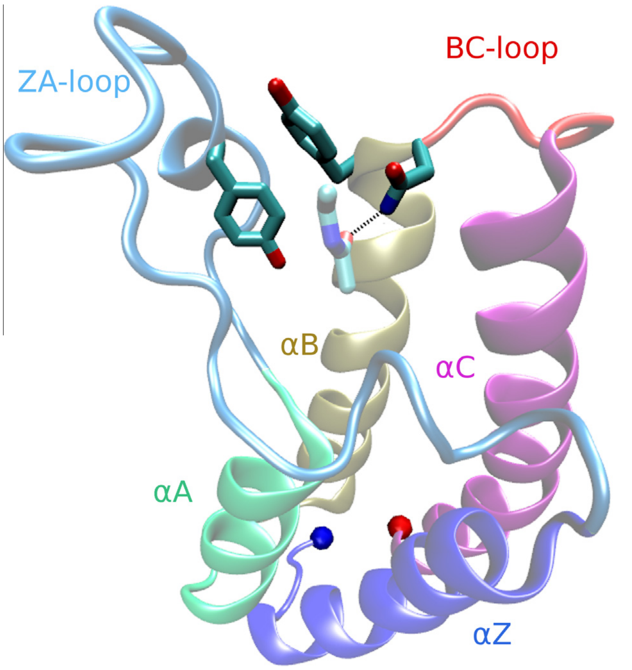


Fig. 1. Ribbon illustration of the crystal structure of the complex between CREBBP and acetyl-lysine (PDB code 3P1C). Each of the four α-helices and two binding site loops is displayed with a different color. The side chains of the conserved Tyr1125 of the ZA-loop, and Tyr1167-Asn1168 of the BC-loop are emphasized (sticks) together with the side chain of acetyl-lysine (sticks, light colors). The N-terminus and C-terminus are shown with a blue and red sphere, respectively.

rotations of the side chain amide of the conserved Asn (Figs. S4 and S5). The rotation (of about 180 degrees) around the χ_2 angle of the conserved Asn switches between an orientation in which the side chain $-NH_2$ group caps the C-terminal turn of α helix B (by donating a hydrogen bond to the backbone carbonyl of the n-4 upstream residue, e.g., the hydrogen bond between the Asn1168 side chain and the carbonyl group of residue 1164 in CREBBP) to a completely solvent exposed orientation of the $-NH_2$ group (Fig. 2, top). Note that the former orientation is observed in most (13 of 17) X-ray structures of the apo bromodomains used in this study (three of them have Thr (PHIP(2) and WDR9(2)) or Tyr (ASH1L) instead of the conserved Asn), as well as in the X-ray structures of BRD4(1) with three different diacetylated peptides (PDB codes 3UVW, 3UVY, and 3UW9) and in the CREBBP/acetyl-lysine complex (PDB code 3P1C). Interestingly, in seven of the 17 bromodomains the Asn $-NH_2$ group is oriented towards the solvent during more than one third of the simulation time (Figs. S4 and S5). These bromodomains are: TAF1L(2) (90% of the snapshots with solvent-exposed orientation of the Asn $-NH_2$ group), TAF1(2) (81%), TAF1(1) (68%), GCN5L2 (44%), ATAD2 (43%), BAZ2B (39%), and PB1(2) (38%). For TAF1L(2), TAF1(2), and TAF1(1) (all belonging to a single clade of family VII in the bromodomain phylogenetic tree [10]) the solvent exposed orientation of the Asn $-NH_2$ group is stabilized by a water-bridged hydrogen bond between the Asn side chain carbonyl and the hydroxyl group of the Ser in position n-4 upstream (Fig. 2, bottom). It has to be noted that most bromodomains have Ala or Cys at position n-4 upstream of the conserved Asn so that these side chains cannot form a (water-bridged) hydrogen bond. Note also that an in-depth analysis of the water molecules in the

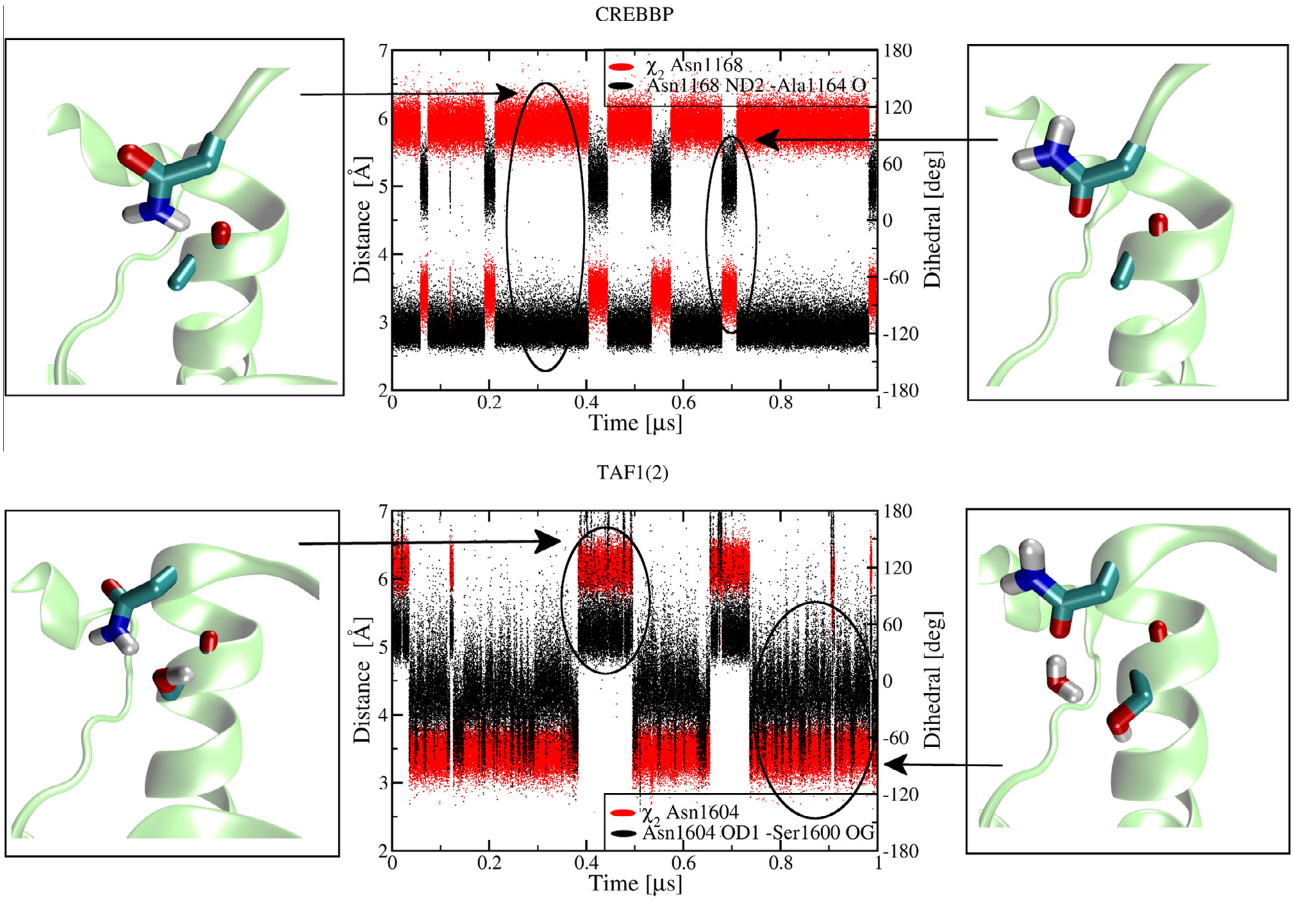


Fig. 2. Rotational flexibility of the conserved Asn side chain as observed in MD simulations of CREBBP (top) and TAF1(2) (bottom). (Middle) Time series of Asn χ_2 angle of conserved Asn (red) and distance between Asn side chain and residue n-4 upstream (black). (Left and right) Representative MD snapshots.

binding site along the MD trajectories goes beyond the scope of this paper and will be presented elsewhere together with a comparison with the five water molecules that seem to be conserved in the crystal structures [10].

We further analyzed the motion of residues adjacent to the conserved Asn to shed light onto the accessibility of the acetyl-lysine binding site. The MD trajectories of some of the 20 simulated bromodomains unmask a large flexibility of the side chain of the evolutionary conserved Tyr (Phe in BAZ2B and SMARCA4) in the BC-loop, i.e., the residue preceding in the sequence the Asn involved in acetyl-lysine binding (Figs. S4 and S5). The side chain of this Tyr (or Phe) rotates from an inward orientation in which the aromatic ring is part of the binding site surface (χ_1 angle in trans) to an outward orientation in which the same side chain points towards the exterior surface of the bromodomain (χ_1 angle values around −60 degrees) (Fig. 3). The five bromodomains with a population larger than 10% of the outward orientation of the BC-loop Tyr (or Phe) side chain are BRDT(1) (100%), CREBBP (31%), BAZ2B (24%), SMARCA4 (21%), and BRD1 (11%). This simulation result is a novel finding because only two of the more than 50 bromodomain crystal and solution structures (i.e., the X-ray structure of BRDT(1) and the NMR structure of KIAA1240, PDB codes 2RFJ and 2DKW, respectively) show a similar outward orientation of the side chain of the conserved Tyr on the BC-loop. Note that in contrast to the solution structure, the crystal structure of KIAA1240 (PDB code 3LXJ) does not show the outward orientation of the conserved Tyr.

To further investigate the influence of the displacement of the BC-loop Tyr on the binding site accessibility, two-dimensional histograms were calculated for each of the 20 simulated bromodomains (Figs. 3(b) and S6). According to the two-dimensional histograms the 20 bromodomains can be assigned to three different groups. The largest group consists of the bromodomains with an open binding site, and in which the side chain of the conserved Tyr of the BC-loop is part of the binding site surface (χ_1 angle always in the trans orientation, e.g., TAF1(2)). The second group includes the bromodomains in which the conserved Tyr changes its orientation without affecting the binding site accessibility (e.g., BRD1 and SMARCA4). The third group includes the bromodomains

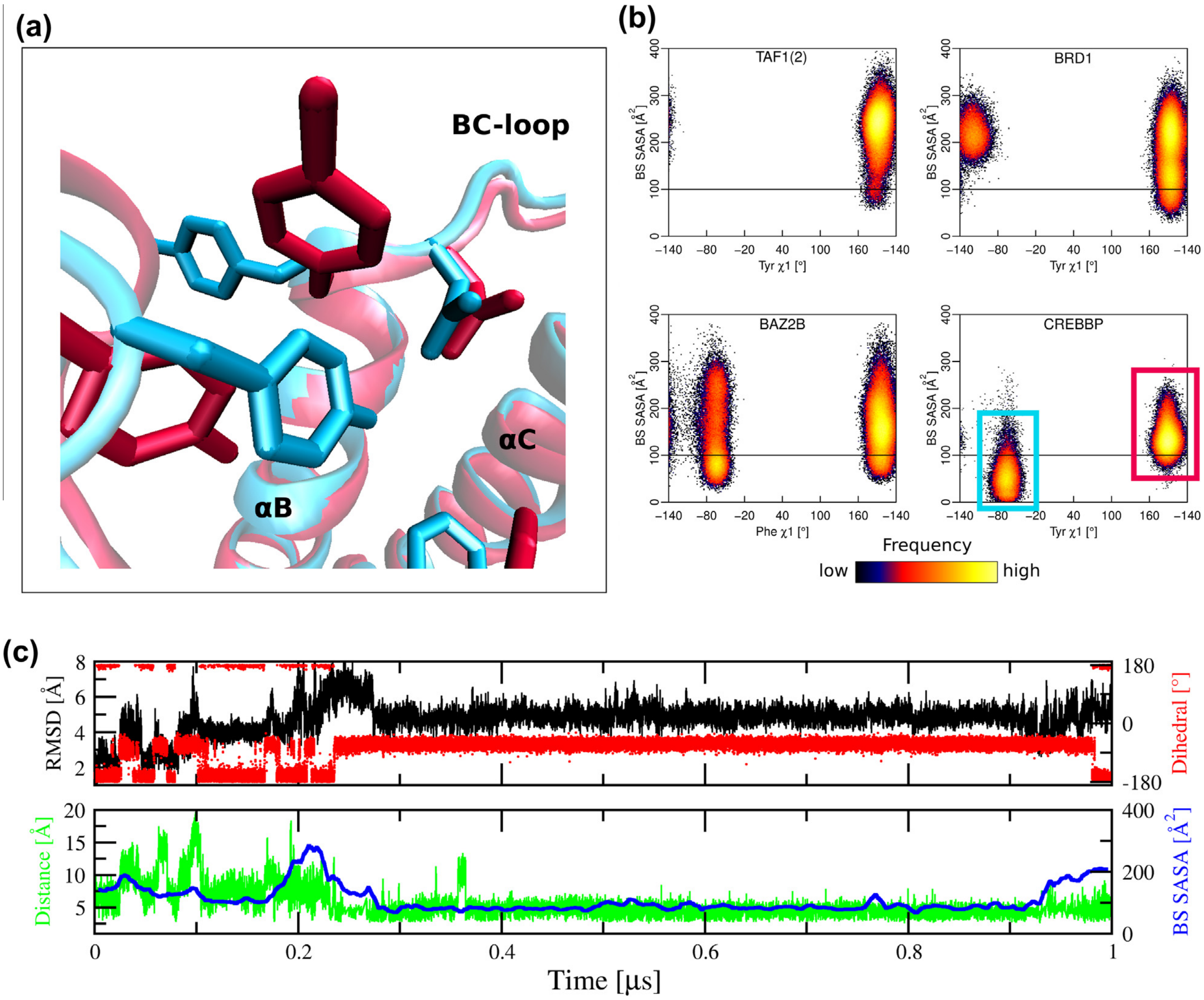


Fig. 3. Swapping out of the aromatic residue preceding the conserved Asn in the BC-loop. (a) Representative snapshots extracted from the MD simulations of CREBBP. (b) Two-dimensional histograms of binding site solvent accessible surface area (SASA) versus side chain χ_1 angle of the conserved Tyr (Phe in BAZ2B) in the BC-loop. The horizontal line at 100 Å^2 is an arbitrary threshold drown solely to help the eye. The cyan and red rectangles in the histogram of CREBBP emphasize the metastable states corresponding to the cyan and red structures, respectively, in panel (a). The black to yellow coloring of frequency values follows a logarithmic scale and is shown in the legend bar. (c) Time series of four geometrical variables used to analyze the binding site. This panel shows run 2 of BAZ2B while the equivalent time series for the other MD runs are shown in Figs. S7–S10. The variables are the RMSD of the C$_\alpha$ atoms of the ZA-loop from the crystal structure (black), the Phe1943 χ_1 dihedral angle (red), the binding site SASA (blue, running average calculated over time intervals of 10 ns for clarity), and the distance between the hydroxyl oxygen of Tyr1901 and the nitrogen atom of the Asn1944 side chain (green).

in which the outward orientation of the BC-loop Tyr promotes a rearrangement of other binding site residues resulting in a partially occluded, i.e., shallower binding site (e.g., BAZ2B and CREBBP, Figs. 3(b) and 4). More precisely, one observes large fluctuations of the ZA-loop that are associated frequently with the swapping out of the BC-loop Tyr (Figs. 3(c), S7 and S8) and with changes in the accessibility of the binding site (Figs. S9 and S10). In particular, in the BAZ2B and CREBBP simulation segments during which the conserved Tyr of the BC-loop points outward, the conformational plasticity of the ZA-loop results (in 20% to 25% of the saved snapshots) in the formation of a direct hydrogen bond between Asn1168 and ZA-loop Tyr1125 in CREBBP (Figs. 3(c) and 4), and Asn1944 and ZA-loop Tyr1901 in BAZ2B. Note that these two side chains are separated by two water molecules in the X-ray structures (PDB codes 3DWY and 3G0L). In essence, it emerges from the MD trajectories that the ZA-loop Tyr side chain can occupy the same position and can be stabilized by the same hydrogen bond to Asn as the natural ligand acetyl-lysine (Fig. 4).

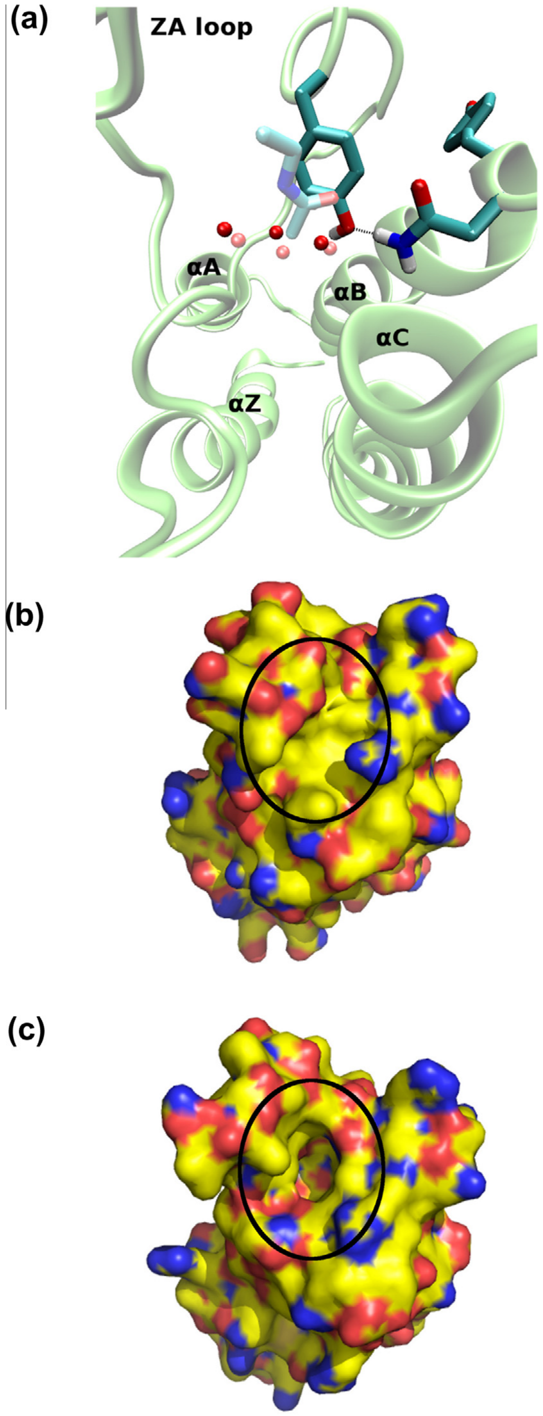


Fig. 4. Occlusion of the acetyl-lysine binding site by the conserved Tyr of the ZA-loop. (a) The ribbon structure is a representative MD snapshot of the CREBBP simulations in which the conserved Tyr of the ZA-loop is involved in a hydrogen bond with the Asn that binds acetyl-lysine. The three conserved side chains (sticks) are shown together with three water molecules (red spheres). The position of acetyl-lysine (light-colored sticks) and three water molecules (light red spheres) from the crystal structure of the CREBBP/acetyl-lysine complex (PDB code 3P1C) was obtained by structural superposition of the C_α atoms. Surface representation of a representative MD snapshot with partially occluded, i.e., shallower, binding site (b) and the crystal structure of CREBBP (PDB code 3P1C) (c). The color coding reflects atomic elements (carbon, nitrogen, and oxygen atoms in yellow, blue, and red, respectively). The location of the acetyl-lysine binding site is indicated with a black oval.

The ZA-loop shows significant plasticity and large deviations from the crystal structure (Figs. 3(c), S7 and S8). This flexibility is consistent with the structural heterogeneity of the ZA-loop observed by overlapping the available X-ray structures [10] as well as with the length of the ZA-loop (about 25–30 residues, Fig. S1). Figs. 3(c), S9 and S10 show that the displacement of the ZA-loop influences the binding site accessibility (note that 4 of the 6 residues used to calculate the binding site accessibility are in the ZA-loop). In particular, the accessibility can change significantly in the MD trajectory segments during which the ZA-loop RMSD from the crystal structure is large, while the accessibility is less variable in segments with mainly rigid ZA-loop. Interestingly, the ZA-loop conformational flexibility has been shown to play an important role at the protein-protein interface for HIV-1 Tat binding to bromodomain PCAF [14]. In the present simulations of CREBBP and BAZ2B, the rearrangement of the ZA-loop and thus of several residues lining the binding site reduces its aperture which could hinder histone peptide recognition.

In conclusion, the MD trajectories reveal a heterogeneous flexibility of the acetyl-lysine binding site in different bromodomains. The analysis of the trajectories shows that the accessibility of the binding site ranges from fully preserved (for the majority of bromodomains) to almost completely inaccessible during simulation segments of hundreds of ns (e.g., in BAZ2B and CREBBP). Of practical importance, the presence of metastable states in a small subset of bromodomains can be exploited [17–19] to design selective small-molecule inhibitors few of which have been reported as of today [4,20–23]. The present simulation results shed some light on the binding mechanism of histone tails to bromodomains which seems to be mainly conformational selection, though a (partial) induced fit cannot be definitively excluded. More precisely, with the present simulations of apo bromodomains it is not possible to exclude a ligand-promoted displacement of some of the side chains in the acetyl-lysine binding site. The self-occluded conformation of the acetyl-lysine binding groove suggests a possible auto-regulatory mechanism of histone tail recognition and binding.

3. Methods

The coordinates of the 20 bromodomains were downloaded from the protein database [24]. To reproduce neutral pH conditions the side chains of aspartates and glutamates were negatively charged, those of lysines and arginines were positively charged, and the histidine side chains were neutral. For each bromodomain the crystal water molecules close to the binding site and with low B factor value were kept while the remaining water molecules were deleted. For most bromodomains a B-factor of 15 Å^2 was used as threshold for keeping crystallographic water molecules. For the few bromodomains with overall higher B-factors a threshold of 30 Å^2 was used. With this arbitrary choice of thresholds between

4 and 15 crystal water molecules were retained for each bromodomain. Importantly, in our simulations the structured water molecules exchange rapidly (residence times ranging from 0.01 to 1 ns) with the water molecules in the bulk solvent. Furthermore, in test simulations without crystal water molecules, bulk water molecules rapidly occupied the conserved positions in the binding site. Thus, the exact number of kept crystal waters does not influence the results of our simulations because individual runs have a length of 0.5 or 1 µs.

Subsequently, the structure was solvated in a water box whose size was chosen to have a minimal distance of 11 Å between the boundary and any atom of the protein. Solvation water molecules within 2.4 Å of any heavy atom of the protein or crystal water were removed. The simulation system contained sodium and chloride ions to approximate an ionic strength of about 150 mM and to compensate for the total charge of each bromodomain. The MD simulations were carried out with Gromacs 4.5.4 [25] using the CHARMM PARAM22 force field [26] and the TIP3P model of water [27].

Periodic boundary conditions were applied and electrostatic interactions were evaluated using the particle-mesh Ewald summation method [28]. The van der Waals interactions were truncated at a cutoff of 10 Å. The temperature of 310 K was kept constant by an external bath with velocity rescaling [29] and the pressure was kept close to 1 atm by the Berendsen barostat. The SHAKE algorithm was used to fix the covalent bonds involving hydrogen atoms. The integration time step was 2 fs and snapshots were saved every 10 ps and 4 ps for the 1-µs and 0.5-µs runs, respectively.

Multiple independent MD runs of 1 µs each were carried out for four bromodomains belonging to four different families: BAZ2B (family V), BRD1 (IV), CREBBP (III), and TAF1(2) (VII). In addition, 16 bromodomains covering all but one of the families were simulated for 0.5 µs each (Table S-I). The crystal structures solved by the Structural Genomics Consortium [10] were used as starting conformations.

The solvent accessible surface area (SASA) was calculated with the SURF tool of Wordom, using the GEPOL algorithm [30,31] and 1.4 Å as radius of the water probe. The following procedure was used to define the acetyl-lysine binding site for the SASA calculation. First, using the X-ray structure of the CREBBP/acetyl-lysine complex (PDB code 3P1C) residues were selected which had at least three heavy atoms within 6 Å of the ligand. From this set Tyr1167 and Asn1168 were excluded to avoid redundancy in the two variables used for the two-dimensional histograms. Of the six remaining residues (Pro1110, Phe1111, Val1115, Tyr1125, Ala1164, Val1174), only the side chain atoms were used to evaluate the binding site SASA. For the other bromodomains the equivalent residues upon structural overlap were chosen. Robustness of the selection of binding site residues is shown in Fig. S11.

Acknowledgments

The simulations were carried out on the Schrödinger cluster at the University of Zurich. This work was supported by a grant of the Swiss National Science Foundation to A.C.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejor.2012.04.007.

References

[1] Zeng, L. and Zhou, M.-M. (2002) Bromodomain: an acetyl-lysine binding domain. FEBS Lett. 513 (1), 124–128.

[2] Filippakopoulos, P. and Knapp, S. (2012) The bromodomain interaction module. FEBS Lett. 586 (17), 2692–2704.

[3] Arrowsmith, C.H., Bountra, C., Fish, P.V., Lee, K. and Schapira, M. (2012) Epigenetic protein families: a new frontier for drug discovery. Nat. Rev. Drug Discov. 11 (5), 384–400.

[4] Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W.B., Fedorov, O., Morse, E.M., Keates, T., Hickman, T.T., Felletar, I., Philpott, M., Munro, S., McKeown, M.R., Wang, Y., Christie, A.L., West, N., Cameron, M.J., Schwartz, B., Heightman, T.D., Thangue, N.L., French, C.A., Wiest, O., Kung, A.L., Knapp, S. and Bradner, J.E. (December 2010) Selective inhibition of bet bromodomains. Nature 468 (7327), 1067–1073.

[5] Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., Taylor, M.J., Johns, C., Chicas, A., Mulloy, J.C., Kogan, S.C., Brown, P., Valent, P., Bradner, J.E., Lowe, S.W. and Vakoc, C.R. (2011) Rnai screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 478 (7370), 524–528.

[6] Dawson, M.A., Kouzarides, T. and Huntly, B.J. (2012) Targeting epigenetic readers in cancer. N. Engl. J. Med. 367 (7), 647–657. PMID: 22894577.

[7] Hewings, D.S., Rooney, T.P.C., Jennings, L.E., Hay, D.A., Schofield, C.J., Brennan, P.E., Knapp, S. and Conway, S.J. (2012) Progress in the development and application of small molecule inhibitors of bromodomain-acetyl-lysine interactions. J. Med. Chem. 55 (22), 9393–9413.

[8] Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. and Zhou, M.-M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. Nature 399 (6735), 491–496.

[9] Mujtaba, S., Zeng, L. and Zhou, M. (2007) Structure and acetyl-lysine recognition of the bromodomain. Oncogene 26 (37), 5521–5527.

[10] Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., Gingras, A.-C., Arrowsmith, C.H. and Knapp, S. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 149 (1), 214–231.

[11] Vidler, L.R., Brown, N., Knapp, S. and Hoelder, S. (2012) Druggability analysis and structural classification of bromodomain acetyl-lysine binding sites. J. Med. Chem. 55 (17), 7346–7359.

[12] Owen, D.J., Ornaghi, P., Yang, J.-C., Lowe, N., Evans, P.R., Ballario, P., Neuhaus, D., Filetici, P. and Travers, A.A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. EMBO J. 19 (22), 6141–6149.

[13] Furdas, S.D., Carlino, L., Sippl, W. and Jung, M. (2012) Inhibition of bromodomain-mediated protein–protein interactions as a novel therapeutic strategy. Med. Chem. Commun. 3, 123–134.

[14] Pantano, S., Marcello, A., Ferrari, A., Gaudiosi, D., Sabò, A., Pellegrini, V., Beltram, F., Giacca, M. and Carloni, P. (2006) Insights on HIV-1 Tat:P/CAF bromodomain molecular recognition from in vivo experiments and molecular dynamics simulations. Proteins: Struct. Funct. Bioinform. 62 (4), 1062–1073.

[15] Pizzitutti, F., Giansanti, A., Ballario, P., Ornaghi, P., Torreri, P., Ciccotti, G. and Filetici, P. (2006) The role of loop ZA and Pro371 in the function of yeast Gcn5p bromodomain revealed through molecular dynamics and experiment. J. Mol. Recognit. 19 (1), :1–9.

[16] Eichenbaum, K.D., Rodríguez, Y., Mezei, M. and Osman, R. (2010) The energetics of the acetylation switch in p53-mediated transcriptional activation. Proteins: Struct. Funct. Bioinform. 78 (2), 447–456.

[17] Ekonomiuk, D., Su, X.-C., Ozawa, K., Bodenreider, C., Lim, S.P., Otting, G., Huang, D. and Caflisch, A. (August 2009) Flaviviral protease inhibitors identified by fragment-based library docking into a structure generated by molecular dynamics. J. Med. Chem. 52 (15), 4860–4868.

[18] Zhao, H., Huang, D. and Caflisch, A. (September 2012) Discovery of tyrosine kinase inhibitors by docking into an inactive kinase conformation generated by molecular dynamics. ChemMedChem 7, 1983–1990.

[19] Laine, E., Martnez, L., Ladant, D., Malliavin, T. and Blondel, A. (August 2012) Molecular motions as a drug target: mechanistic simulations of anthrax toxin edema factor function led to the discovery of novel allosteric inhibitors. Toxins (Basel) 4 (8), 580–604.

[20] Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C.-W., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., White, J., Kirilovsky, J., Rice, C.M., Lora, J.M., Prinjha, R.K., Lee, K. and Tarakhovsky, A. (December 2010) Suppression of inflammation by a synthetic histone mimic. Nature 468 (7327), 1119–1123.

[21] Gerona-Navarro, G., Yoel-Rodrguez, Mujtaba, S., Frasca, A., Patel, J., Zeng, L., Plotnikov, A.N., Osman, R. and Zhou, M.-M. (February 2011) Rational design of cyclic peptide modulators of the transcriptional coactivator CBP: a new class of p53 inhibitors. J. Am. Chem. Soc. 133 (7), :2040–2043.

[22] Chung, C.-W., Dean, A.W., Woolven, J.M. and Bamborough, P. (January 2012) Fragment-based discovery of bromodomain inhibitors part 1: Inhibitor binding modes and implications for lead discovery. J. Med. Chem. 55 (2), 576–586.

[23] Bamborough, P., Diallo, H., Goodacre, J.D., Gordon, L., Lewis, A., Seal, J.T., Wilson, D.M., Woodrow, M.D. and Chung, C.-W. (January 2012) Fragment-based discovery of bromodomain inhibitors part 2: Optimization of phenylisoxazole sulfonamides. J. Med. Chem. 55 (2), 587–596.

[24] Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1978) The protein data

bank: a computer-based archival file for macromolecular structures. Arch. Biochem. Biophys. 185 (2), 584–591.

[25] Spoel, D.V.D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J.C. (December 2005) Gromacs: fast, flexible, and free. J. Comput. Chem. 26 (16), 1701–1718.

[26] MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L.J., et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 102, 35863616.

[27] Jorgensen, W.L., Chandrasekhar, J., Madura, J., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926935.

[28] Darden, T., York, D. and Pedersen, L.G. (1993) Particle mesh Ewald: an $n\log(n)$ method for Ewald sums in large systems. J. Chem. Phys. 98, 10089.

[29] Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical sampling through velocity rescaling. J. Chem. Phys. 126 (1), 014101.

[30] Seeber, M., Cecchini, M., Rao, F., Settanni, G. and Caflisch, A. (October 2007) Wordom: a program for efficient analysis of molecular dynamics simulations. Bioinformatics 23 (19), 2625–2627.

[31] Seeber, M., Felline, A., Raimondi, F., Muff, S., Friedman, R., Rao, F., Caflisch, A. and Fanelli, F. (April 2011) Wordom: a user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. J. Comput. Chem. 32 (6), 1183–1194.

Does bromodomain flexibility influence histone recognition?

SUPPLEMENTARY MATERIAL

S. Steiner[a], A. Magno[a], D. Huang[a], and A. Caflisch[a]*

[a]Department of Biochemistry
University of Zürich, Winterthurerstrasse 190
CH-8057 Zürich, Switzerland
Phone: (+41 44) 635 55 21, FAX: (+41 44) 635 68 62
email: caflisch@bioc.uzh.ch

*Corresponding author

keywords: molecular dynamics, post-translational modifications, histones, epigenetic targets

May 9, 2013

Simulation systems

Bromodomain	Family	PDB entry	Number of independent runs	Total sampling (μs)	Number of available X-ray/NMR structures
CECR2	I	3NXB	1	0.5	1
FALZ	I	3UV2	1	0.5	2
GCN5L2	I	3D7C	1	0.5	2
PCAF	I	3GG3	1	0.5	2
BRD4(1)	II	2OSS	1	0.5	1
BRDT(1)	II	2RFJ	1	0.5	1
BRD3(2)	II	2OO1	1	0.5	2
PHIP(2)	III	3MB3	1	0.5	1
WDR9(2)	III	3Q2E	1	0.5	1
CREBBP	III	3DWY	5	**4.5**	1
EP300	III	3I3J	1	0.5	1
ATAD2	IV	3DAI	1	0.5	1
BRD1	IV	3RCW	5	**4.5**	1
BAZ2B	V	3G0L	5	**4.5**	1
TAF1(1)	VII	3UV5	1	0.5	2
TAF1L(2)	VII	3HMH	1	0.5	1
TAF1(2)	VII	3UV4	4	**4.0**	2
ASH1L	VIII	3MQM	1	0.5	1
PB1(2)	VIII	3HMF	1	0.5	1
SMARCA4	VIII	3UVD	1	0.5	3

Table S-I: The digit in parentheses in the first column indicates the position of the bromodomain along the polypeptide sequence, e.g., TAF1(2) is the second bromodomain of the human Transcription initiation factor TFIID subunit 1 (TAF1). Multiple independent MD runs were carried out for the four bromodomains in boldface. The family classification and number of available X-ray/NMR structures are taken from Ref.[1]

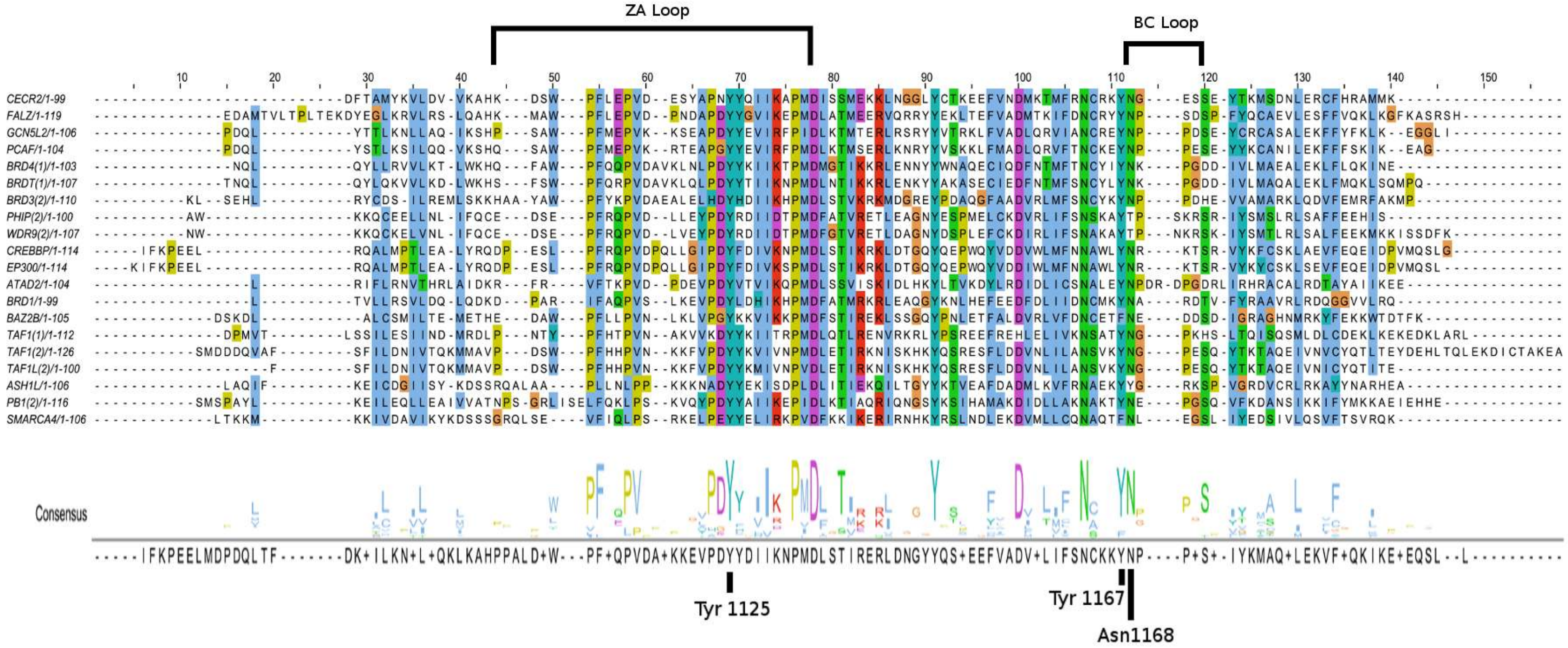


Figure S1: **Sequence alignment of the 20 simulated bromodomains.** The sequences were aligned first with MUSCLE ,[2] and subsequently realigned with ClustalW[3] using default setting parameters. Residues are colored following the ClustalX color scheme. The three residues in the bottom emphasize the evolutionary conserved Tyr side chain on the ZA-loop, i.e., Tyr1125 in CREBBP, Tyr599 in BRD1, Tyr1901 in BAZ2B, and Tyr1561 in TAF1(2), and the conserved aromatic residue and Asn at the beginning of the BC-loop, i.e., Tyr1167-Asn1168 in CREBBP, Tyr641-Asn642 in BRD1, Phe1943-Asn1944 in BAZ2B, and Tyr1603-Asn1604 in TAF1(2). The conserved residues are labeled only for CREBBP.

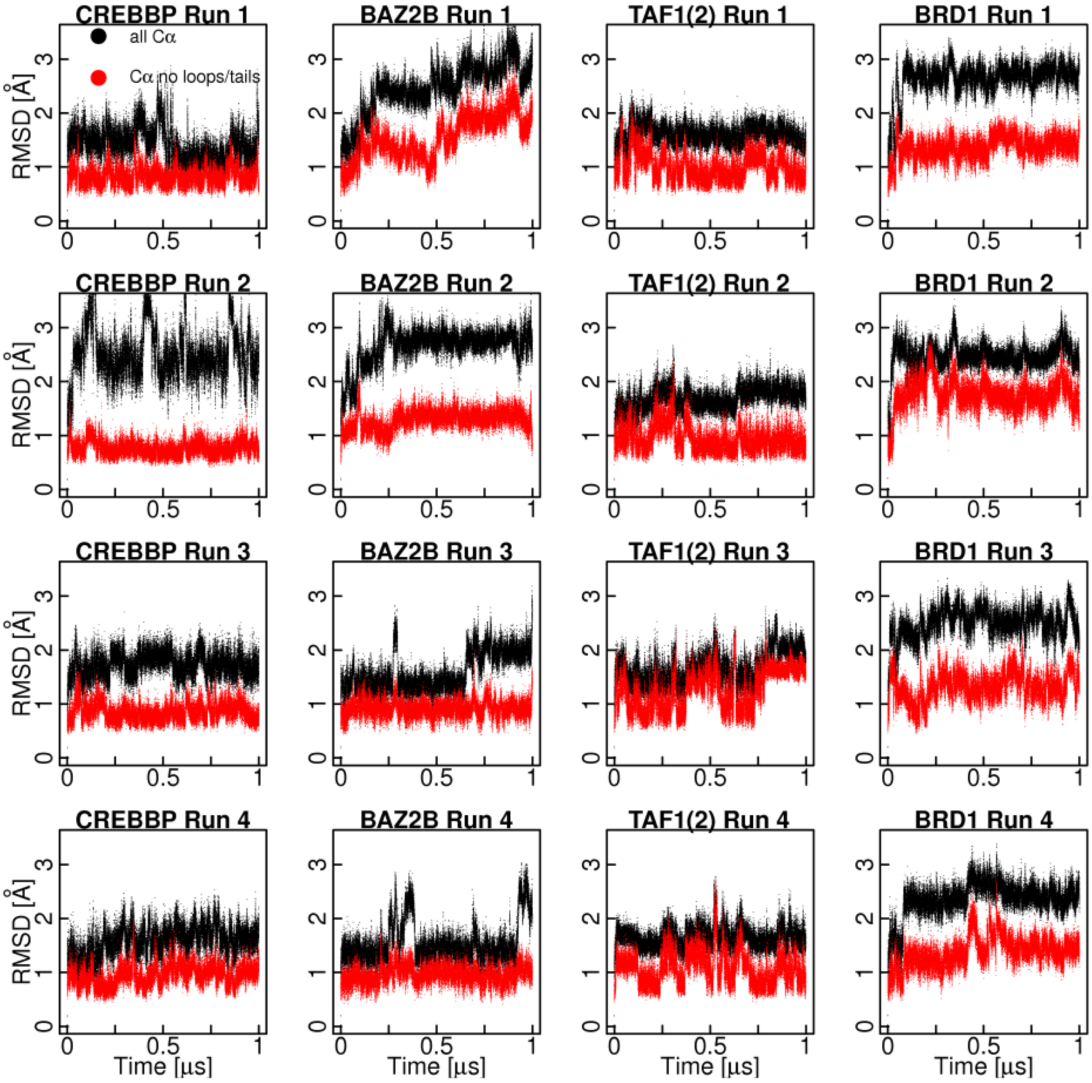


Figure S2: **Root mean square deviation (RMSD) of C_α atoms.** The time series of the RMSD from the minimized structures (i.e., starting structure of the MD runs) is plotted for all C_α atoms (black line) and excluding the C_α atoms in the ZA-loop, BC-loop, and the termini (red line). For CREBBP C_α's of residues 1104-1134 and 1168-1172 as well as 3 terminal residues on each N- and C-terminus were excluded for the latter calculation. For BAZ2B, BRD1 and TAF1(2) the corresponding residues upon structural overlap were neglected. Note that similar RMSD values are observed for the other 16 bromodomains.

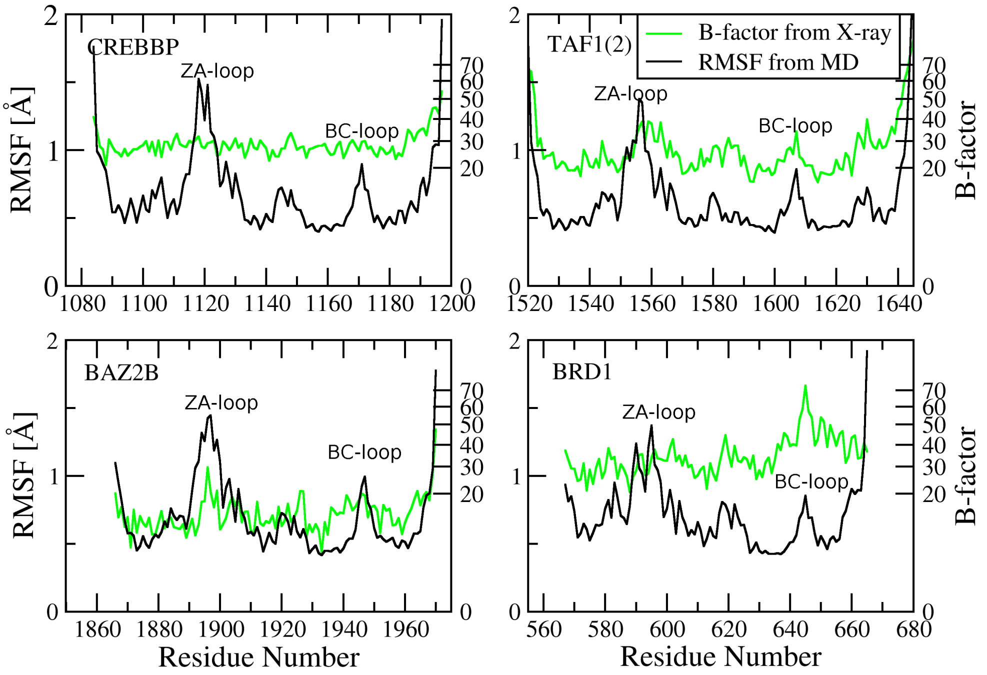


Figure S3: **Root mean square fluctuations (RMSF).** Profile of the RMSF of the C_α atoms (black line, y-axis on the left) of CREBBP, BAZ2B, BRD1 and TAF1(2) and crystallographic B-factors (green line, y-axis on the right) as a function of residue number. The RMSF values were averaged over simulation intervals of 5 ns. The fluctuations are highest for the ZA-loop, BC-loop, and termini. Note that the major discrepancies between RMSF along the simulations and crystallographic B-factors are due to crystal contacts in the latter, e.g., for the loops of CREBBP which are stabilized by the nearest neigbors in the crystal (PDB code 3DWY).

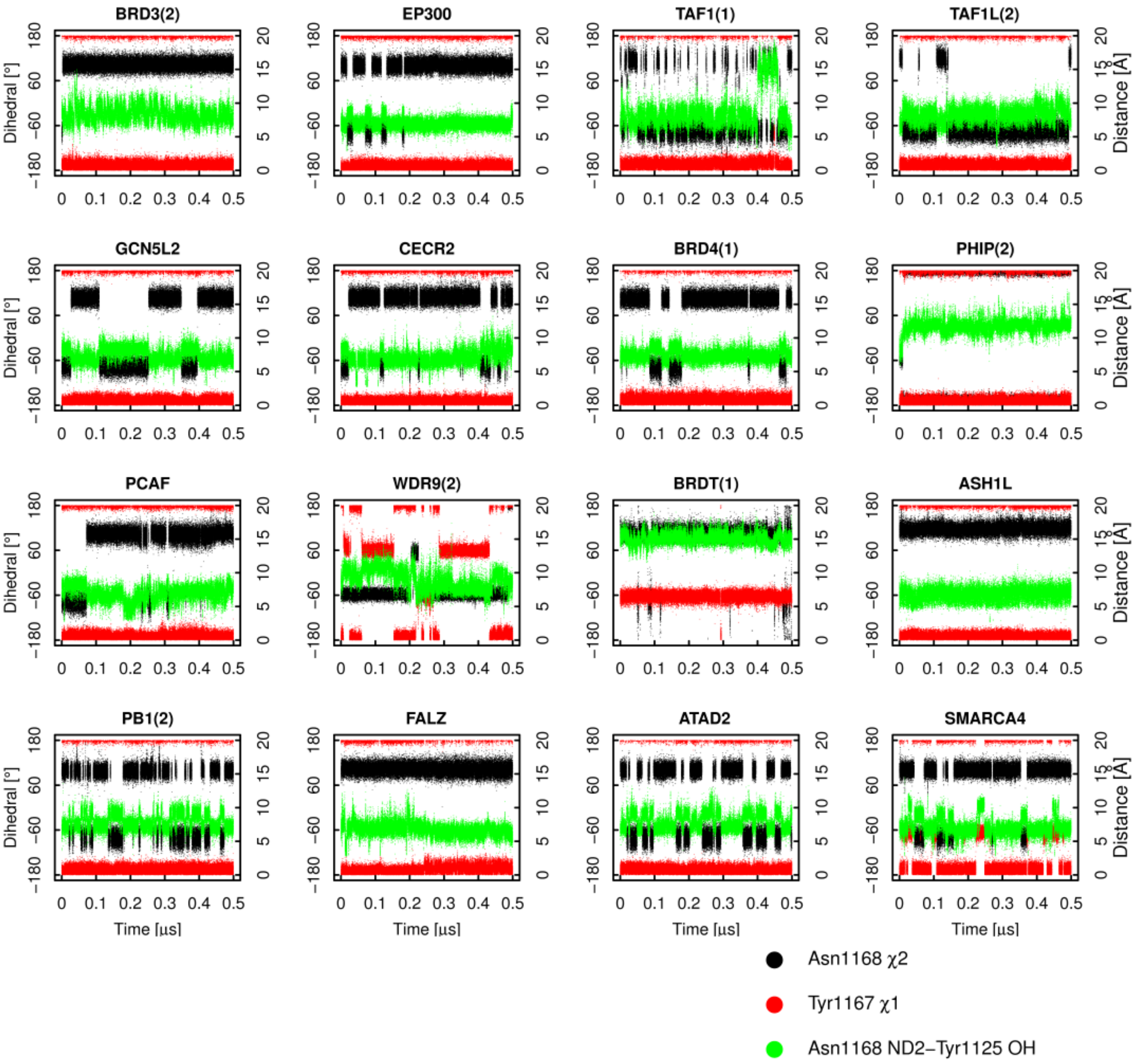


Figure S4: **Conserved Asn and Tyr/Phe: Time series of dihedrals and distances.** The χ_2 dihedral of the conserved Asn (i.e., Asn1168 of CREBBP) is shown (black). Note that for PHIP(2) and WDR9(2) χ_1 of Thr equivalent to Asn1168, and for ASH1L χ_2 of Tyr equivalent to Asn1168 are shown. The χ_1 of Tyr/Phe preceding in sequence the conserved Asn (i.e., Tyr1167 in CREBBP) is also shown (red). The distance between the side chain ND2 of the conserved Asn and the hydroxyl oxygen of the ZA-loop Tyr pointing inside the binding groove (i.e., Tyr1125 in CREBBP) is also shown (green); in case of a Thr or Tyr in position of the conserved Asn, Thr OG1, resp. Tyr OH were used instead of Asn ND2. Note that the ZA-loop Tyr is conserved in all the 20 bromodomains investigated here (see Figure S1).

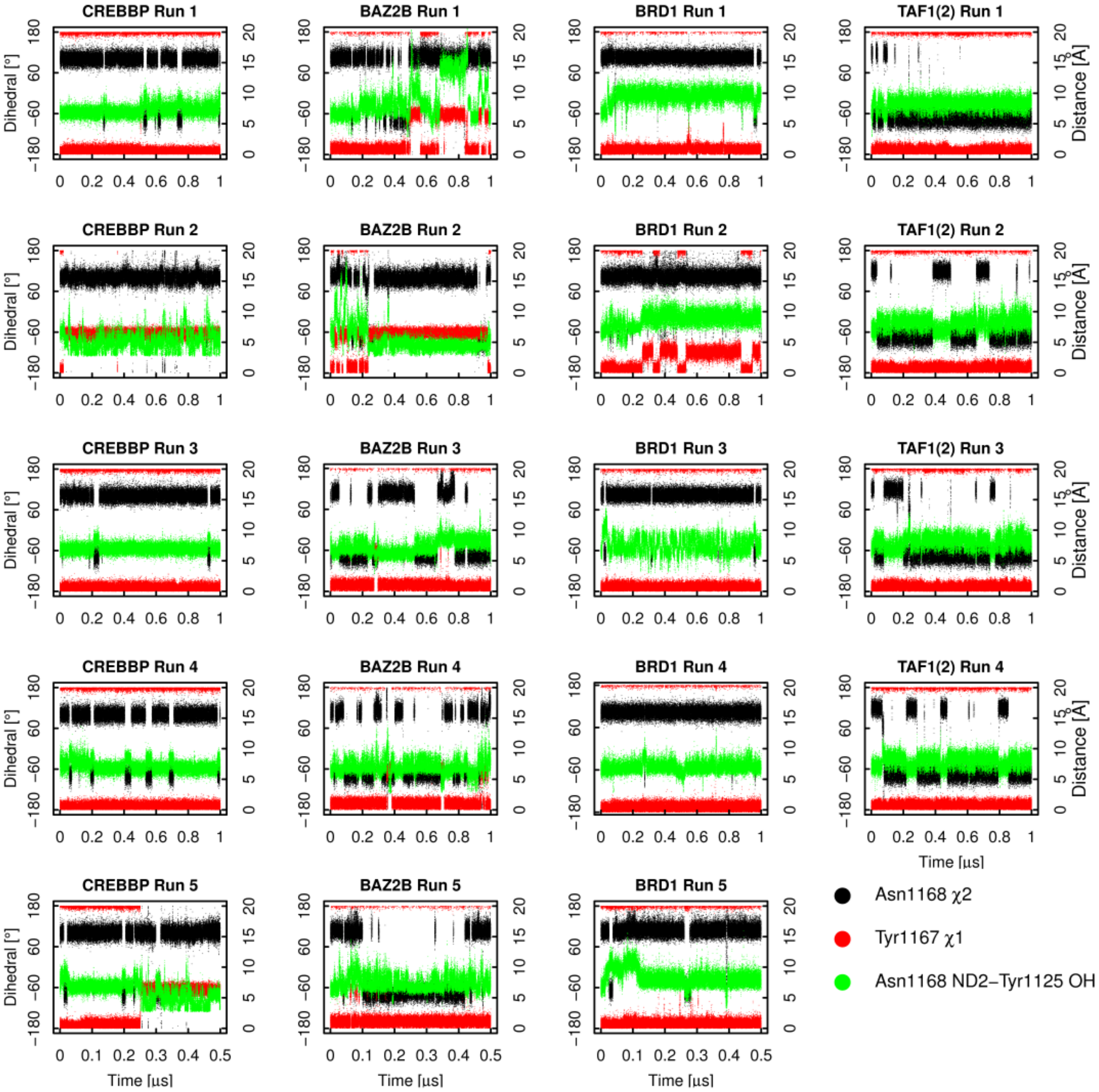


Figure S5: **Conserved Asn and Tyr/Phe: Time series of dihedrals and distances.** Same as Figure S4 for the four bromodomains for which multiple MD simulations were carried out.

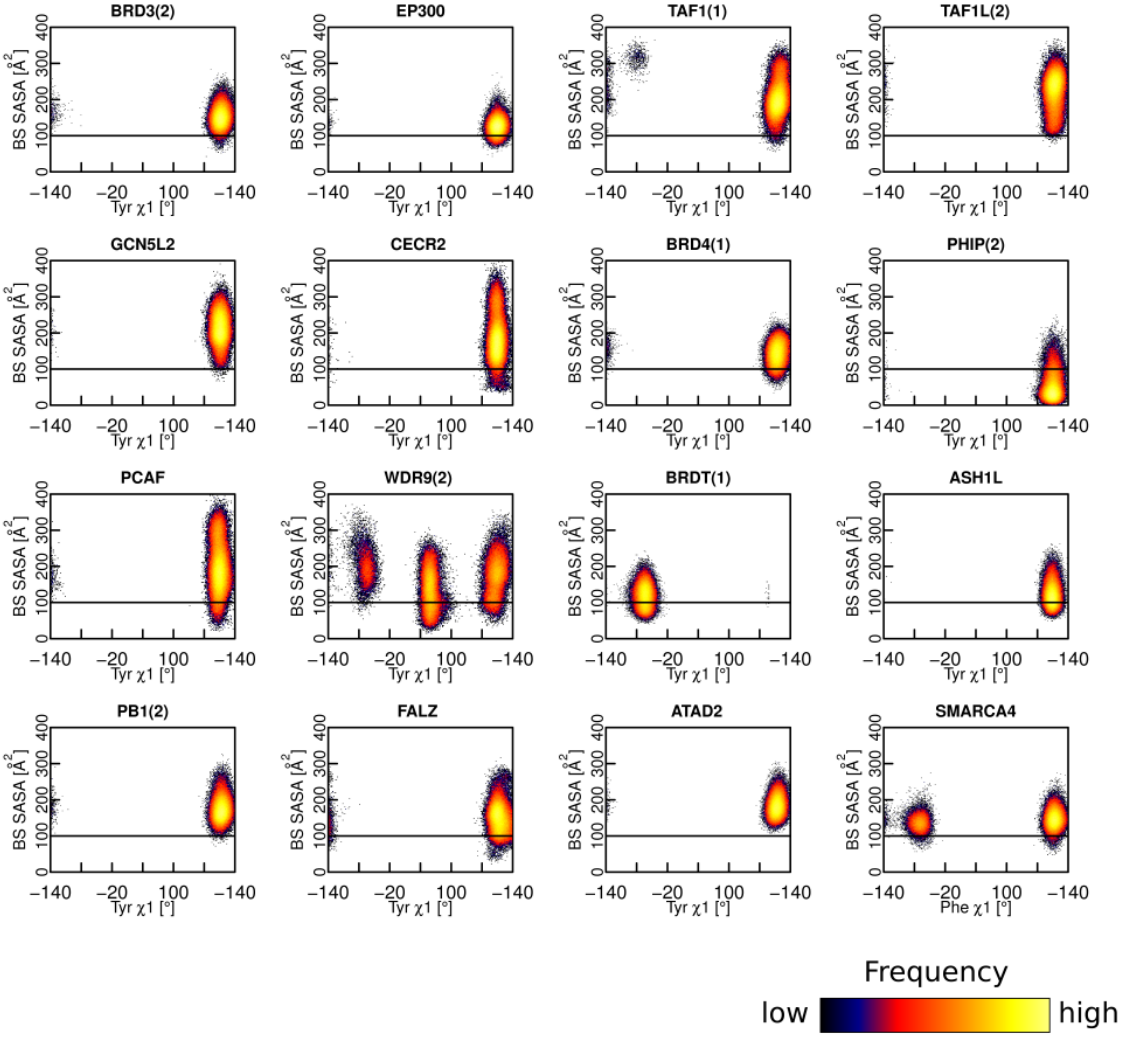


Figure S6: **Binding site accessible surface.** Two-dimensional histograms of binding site accessible molecular surface (SASA) versus χ_1 dihedral of Tyr/Phe preceding the conserved Asn in the BC-loop. Coloring according to frequencies follows a logarithmic scale.

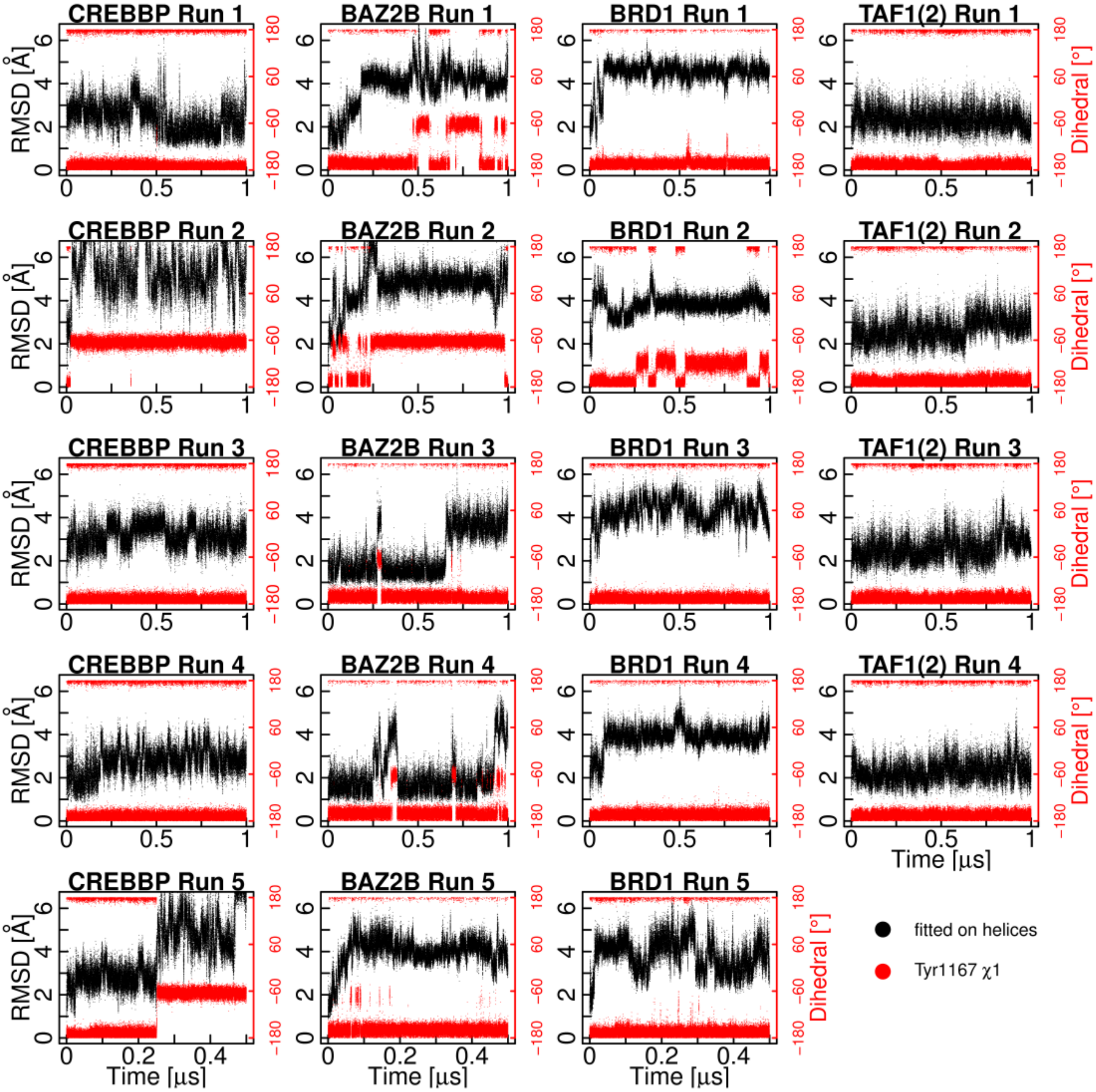


Figure S7: **Flexibility of ZA-loop and side chain rotation of the conserved Tyr/Phe in BC-loop.** Each panel corresponds to an independent simulation (of the bromodomain indicated in the top) and shows the time series of RMSD of the C_α atoms of the ZA-loop (black, y-axis on the left) and the χ_1 dihedral angle of the conserved Tyr/Phe in the BC-loop (red, y-axis on the right). The RMSD of the ZA-loop from the energy-minimized crystal structure was calculated upon structural overlap of the MD snapshots using the C_α atoms of the four α-helices.

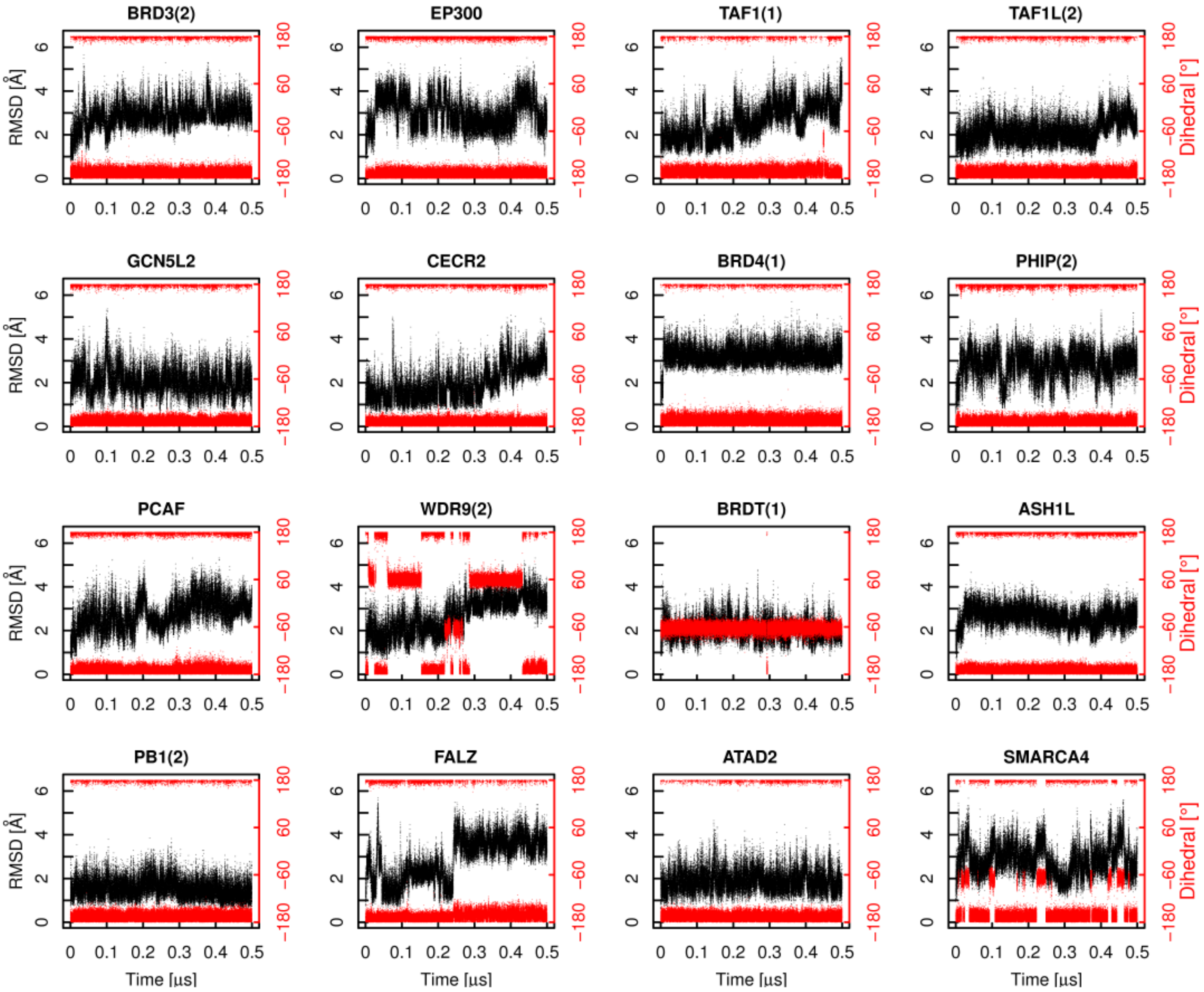


Figure S8: **Flexibility of ZA-loop and side chain rotation of the conserved Tyr/Phe in BC-loop.** See caption of Figure S7.

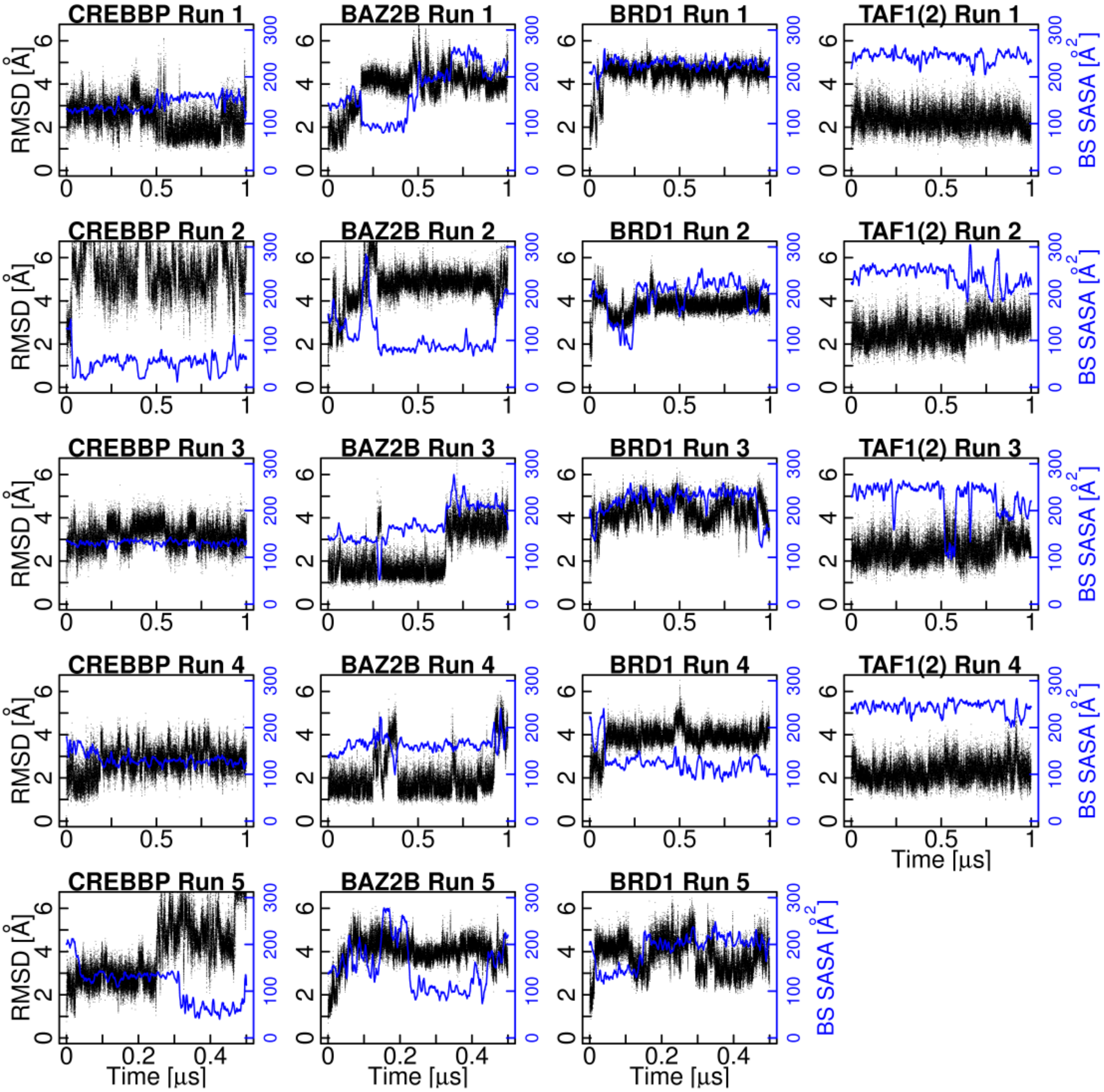


Figure S9: **Flexibility of ZA-loop and changes in binding site SASA.** Each panel corresponds to an independent simulation (of the bromodomain indicated in the top) and shows the time series of RMSD of the C_α atoms of the ZA-loop (black, y-axis on the left) and the binding site SASA (blue, y-axis on the right, running average calculated over time intervals of 10 ns for clarity). The RMSD of the ZA-loop from the energy-minimized crystal structure was calculated upon structural overlap of the MD snapshots using the C_α atoms of the four α-helices.

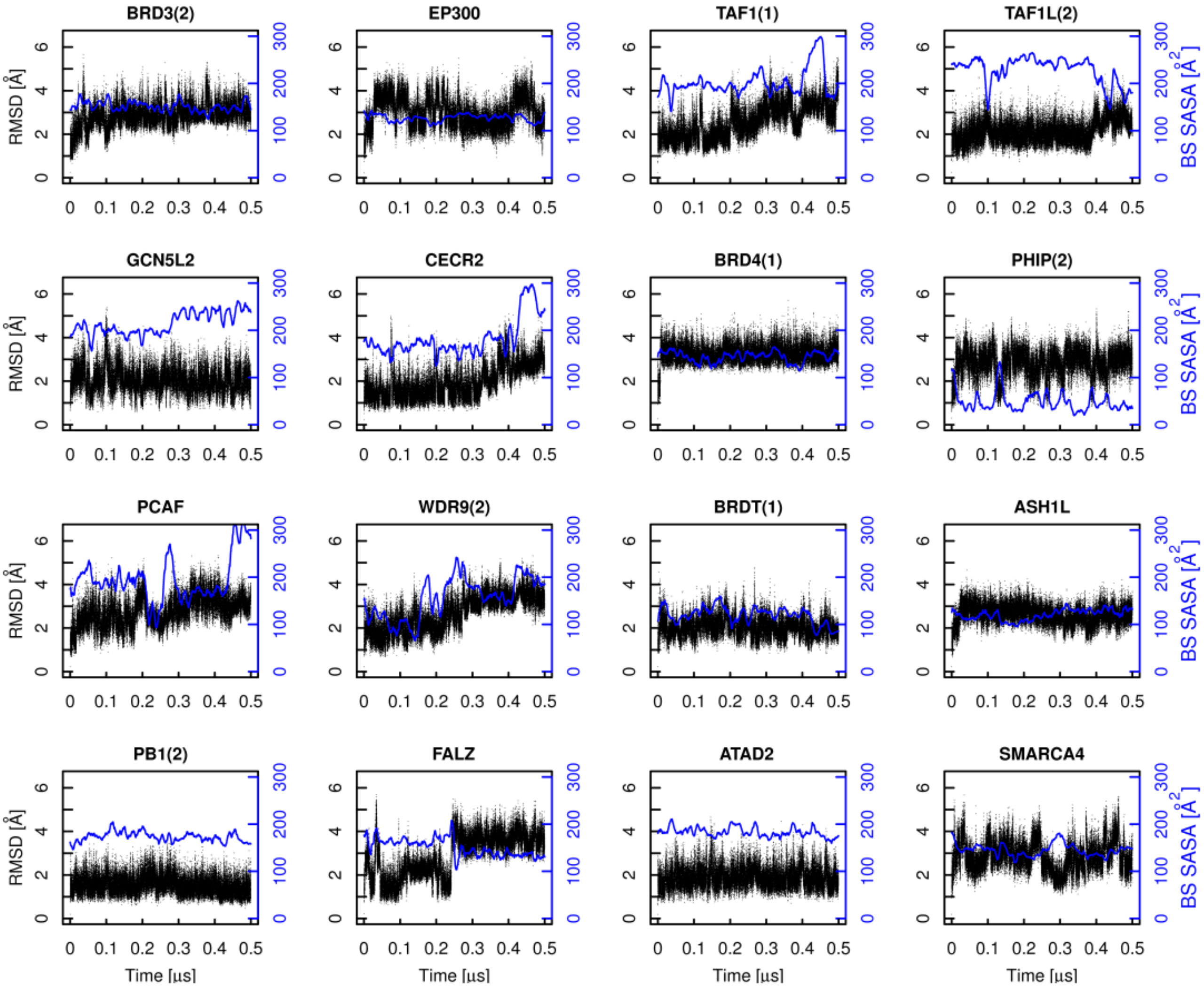


Figure S10: **Flexibility of ZA-loop and changes in binding site SASA.** See caption of Figure S9.

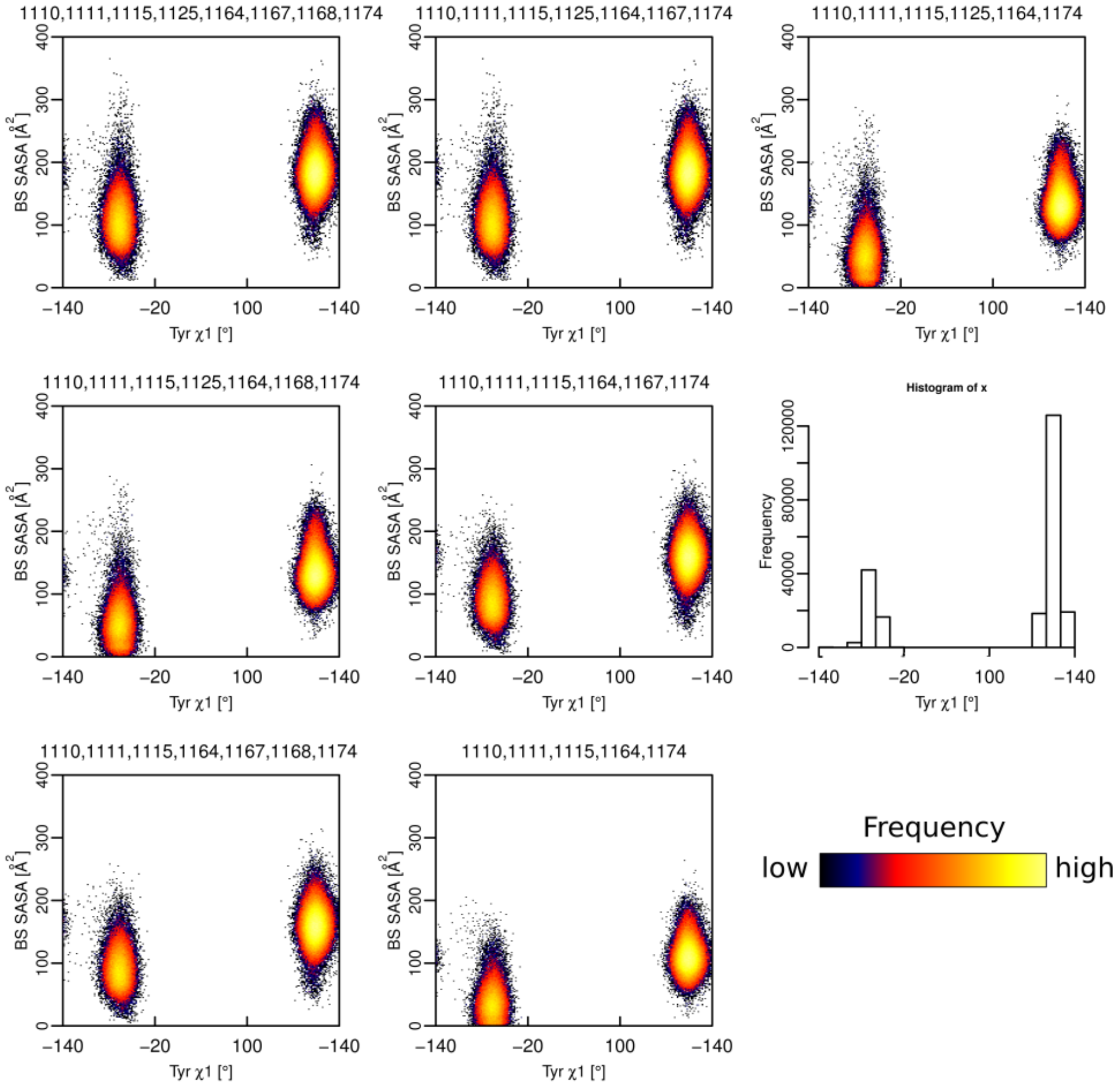


Figure S11: **Robustness of two-dimensional histograms.** The very similar two-dimensional histograms of binding site SASA vs. χ_1 of Tyr1167 for CREBBP, calculated using seven different selections of binding site atoms, demonstrate robustness. The subset of binding site residues used for each of the seven panels are the side chain heavy atoms of the residues whose numbers are mentionend above each panel. Coloring according to frequencies follows a logarithmic scale. The one-dimensional histogram of the χ_1 of CREBBP Tyr1167 is also shown (middle right).

References

[1] Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., Gingras, A.-C., Arrowsmith, C. H., and Knapp, S. Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 149(1):214–231, 2012.

[2] Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5):1792–1797, 2004.

[3] Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Thompson, J., Gibson, T., and Higgins, D. Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948, 2007.

Chapter 8

Conclusions

Molecular dynamics (MD) simulations are a powerful tool, which are widely employed in several scientific areas since they provide a direct route from the microscopic details of a system to its macroscopic, experimentally observable features. In this thesis, a wide spectrum of MD techniques were employed to study several different biophysical issues, which range from peptides self-assembly in a crowded environment or in presence of a small inhibitor, to bromodomains recognition of acetylation marks. The numerous molecular processes investigated in this thesis differ in their characteristic time scales and in the complexity of the microscopic mechanisms they involve.

Using a coarse-grained phenomenological model, we could assess how macromolecular crowding affects aggregation kinetics, showing that its net effect on the self-assembly process depends crucially on the intrinsic amyloidogenic propensity of the studied peptide: the less a peptide is prone to aggregate, the greater is the acceleration exerted on its aggregation from macromolecular crowding.

Implicit solvent constant temperature MD simulations were performed to shed light on the structures of the early aggregates of the amino acid phenylalanine, whose fibrillar assemblies were shown in experiments to have a cytotoxic effect and were then hypothesized to play a crucial role in the etiology of Phenylketonuria. The structure we proposed was in very good agreement with the electron diffraction pattern experimentally obtained.

The use of replica exchange MD simulations in conjunction with an implicit solvent model allowed us to study the interaction of carnosine with the amyloid β-peptide (Aβ) and to explain its role as an amyloid aggregation modulator. In agreements with experimental studies, our results showed that carnosine establishes transient, non-specific interactions with the Aβ monomer and that its conformational ensemble is not significantly affected from the presence of carnosine. However, a reduction in the formation of interpeptide hydrogen bonds in presence of carnosine was observed in simulations of Aβ peptide segments.

Explicit solvent MD simulations were performed to investigate the binding site plasticity of 20 bromodomains and the binding process of acetylated lysine to one of them. Simulations revealed a large variability in the available orientation of conserved residues in the binding cleft of several bromodomains, which for some of them was also accompanied by partial occlusion of their binding site. Moreover, we demonstrated that our computational approach is able to reproduce the experimentally observed binding pose of acetyl-lysine to various bromodomains. Interestingly, our study unveiled the presence of a second stable binding pose, and showed that the system can interconvert between the two binding conformations on a much faster time scale with respect to the binding/unbinding kinetics.

The results collected in this thesis prove that molecular dynamics techniques are indeed a versatile and useful investigation method in biophysics, which is able to corroborate and extend the picture of molecular processes obtained by experiments and theoretical models.

Acknowledgments

It seems to me to have just arrived yesterday here in Zurich to be interviewed by Prof. Amedeo Caflisch, and I am already writing the Acknowledgments of my PhD thesis. The doctoral studies have brought me far from my beautiful Rome and the books full of equations I used to read as a student of Physics. However, they have given me a new home in the bright Zurich and the opportunity to study the field of Biochemistry. This would not have been possible, if on this day far in the past in September Prof. Amedeo Caflisch had not given me the chance to join his group as a PhD student. He is the first person I have to thank for all that has happened in the last almost 4 years. I enjoyed working with him, his positive and liberal leadership has helped me several times during my studies. I have also another important reason to thank him, that is, if he had not hired me here, I would never have met Sandra, the second person I want to thank here, for her kind friendship first, the scientific collaboration then, and, finally, for being the dear and beloved partner she is now for me. A special thank goes to Riccardo Pellarin, who took me under his wing when I had just arrived here in Zurich and with whom I had the pleasure to work during the first year of my doctoral studies. Then come all the extraordinary people I have met here, who have accompanied me in this time: Fabio e Davide, che sono stati per me come una famiglia qui a Zurigo, Lorenzo e Daniela, per le danze sfrenate e le lunghe, lunghissime chiacchierate, Beatrice e Andrea S., per tutto l'aiuto, la generosità e i bei pranzi domenicali a Schwamendingen, Michela e Francesca, per le divertenti serate trascorse insieme, Andrea P., per le chiacchierate alle 3 di notte nella cucina di Mutschellenstrasse 89, Chiara and Benoît, with whom I went through the first summer for me here in Zurich, I have explored California, Sicily and even Milan and the alpin landscape around Varese, John, who has showed me the sweet hills around Vevey and has brought me to the Jagger in Lausanne, and Pietro, Enrico, Riccardo S. and Nicolas, not just working colleagues but dear friends to me. E poi quelli che mi sono stati sempre vicini, nonostante la lontananza: mamma e papà, per tutto il supporto e l'affetto che mi avete dato nonostante i quasi 1000 km di distanza, i miei fratelli Lorenzo ed Edoardo, che si sono quasi svenati per venirmi a trovare qui a Zurigo, Valerio, l'amico di una vita, Francesca, l'amica di una vita, Andrea e Francesca, perché anche se ci vediamo solo tre volte l'anno quando va bene è sempre una gioia. E ovviamente voi, Alessia, Anna, Alessandra, Maria e Renata, per il bene immenso che ci vogliamo e perché ogni volta che ci rivediamo è una festa. Finally I would like to thank all the colleagues I had the pleasure to work with in these 4 years, especially Sandra R. and Emilie, for their support and kindness. E poi avrei voluto ringraziare te François, carissimo amico prima che generoso e insostituibile collega, sempre pronto ad aiutare tutti senza mai chiedere niente in cambio. Ma te ne sei andato ormai quasi un anno fa', lasciando tra noi un vuoto incolmabile, in questo gruppo, in questo ufficio e nelle nostre vite. Che dire, è proprio vero che ad andarsene son sempre i migliori.

List of publications

Crowding effects on amyloid aggregation kinetics.
Magno, A., Caflisch, A. and Pellarin, R.
[Jour. Phys. Chem. Lett., 2010, 1, 3027-3032.]

Mechanisms and kinetics of amyloid aggregation investigated by a phenomenological coarse-grained model.
Magno, A., Pellarin, R. and Caflisch, A.
[Chapter of the book: "Computational Modeling of Biological Systems: From Molecules to Pathways", 2012, 191-214, Springer.]

Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria.
Adler-Abramovich, L., Vaks, L., Carny, O., Trudler, D., Magno, A., Caflisch, A., Frenkel, D. and Gazit, E.
[Nat. Chem. Biol., 2012, 8, 701-706.]

Carnosine inhibits $A\beta_{42}$ aggregation by perturbing the H-bond network in and around the central hydrophobic cluster.
Attanasio, F., Convertino, M., Magno, A., Caflisch, A., Corazza, A., Haridas, H., Esposito, G., Cataldo, S., Pignataro, B., Milardi, D. and Rizzarelli, E.
[ChemBioChem, 2013, 14, 583-592.]

Mechanism and kinetics of acetyl-lysine binding to bromodomains.
Magno, A.[‡], Steiner, S.[‡] and Caflisch, A.
[J. Chem. Theory Comput., 2013, 9, 4225-4232.]
[‡]These authors contributed equally to this work

Does bromodomain flexibility influence histone recognition?
Steiner, S., Magno, A., Huang, D. and Caflisch, A.
[FEBS Lett., 2013, 587, 2158-2163.]

Curriculum Vitae

Andrea Magno

born 27th February 1985 in Rome, Italy

Italian citizen

Education

November 2009 - July 2013	PhD student in the group of Prof. Dr. Amedeo Caflisch, Department of Biochemistry, University of Zurich.
July 2009	Master degree in Physics at the University of Roma Tre, score 110/110 *cum laude*. Title of the masterwork: “Molecular Dynamics Simulations for the Study of the Slow Dynamics and of the Strong-Fragile Transition of Supercooled Aqueous Solutions of Trehalose”.
September 2004 - July 2009	Studies of Physics at the University of Roma Tre.
September 1999 - July 2004	High School Diploma at the Liceo Classico Statale “Socrate”, Rome.